VIX derivatives, hedging and vol-of-vol risk

Kaeck, Andreas; Seeger, Norman J.

published in
European Journal of Operational Research
2020

DOI (link to publisher)
10.1016/j.ejor.2019.11.034

document version
Publisher's PDF, also known as Version of record

document license
Article 25fa Dutch Copyright Act

European Journal of Operational Research 283 (2020) 767–782

Contents lists available at ScienceDirect

European Journal of Operational Research

journal homepage: www.elsevier.com/locate/ejor

Interfaces with Other Disciplines

VIX derivatives, hedging and vol-of-vol risk

Andreas Kaeck[a], Norman J. Seeger[b,*]

[a] University of Sussex Business School, Falmer, Brighton BN1 9RH, United Kingdom
[b] VU University Amsterdam, The Netherlands

ARTICLE INFO

Article history:
Received 19 May 2018
Accepted 17 November 2019
Available online 21 November 2019

Keywords:
Risk management
VIX options
Stochastic volatility of volatility
Hedging performance,

ABSTRACT

We study the empirical hedging performance of alternative VIX option pricing models. Recent advances in the literature find evidence of asymmetric volatility-of-volatility (similar to the leverage effect in equity markets), stochastic mean-reversion and jumps. Using such findings in our model framework, we show that while sophisticated models have superior pricing performance and can explain a range of stylized facts in the VIX derivatives market, their hedging performance is inferior to a simple Black model hedge. We also study the empirical performance of regime-dependent hedge ratio adjustments commonly applied in equity markets.

1. Introduction

This paper provides a comprehensive empirical study of the hedging performance of alternative models for Volatility Index (VIX) derivatives. To date the operations research literature studying derivative markets has focused predominantly on option pricing questions (see Bandi & Bertsimas, 2014; Broadie & Detemple, 2004; Fu, Li, Li, & Wu, 2016; Liu, Cao, Ma, & Shen, 2019; Wong & Lo, 2009 and others). We contribute to this literature by developing pricing models for VIX derivatives, a fairly new asset class that has seen a large increase in trading volume over the past years. Rather than focussing on assessing models using pricing performance metrics, our methodology takes a risk management perspective by devising empirical out-of-sample tests based on the performance of alternative hedging models. This allows to circumvent some of the criticism in the literature that argues that dynamic hedging strategies may be more suitable for testing alternative model specifications than pricing error metrics.

The VIX index is considered a barometer of investor sentiment and market-wide volatility.[1] It has also been shown that the index is informative about future economic activity and that shocks to the VIX index affect real economic activity (see Berger, Dew-Becker, & Giglio, 2019; Bloom, 2009). Since the VIX index is not a directly tradeable instrument, the Chicago Board Options Exchange (CBOE) introduced VIX derivatives to allow investors to create tradeable exposure to volatility. VIX futures and options are widely used in risk management strategies and present two of the most successful product launches of the CBOE.[2] VIX derivatives are also building blocks for other traded instruments, Alexander and Korovilas (2013) study a range of VIX-related Exchange Traded Notes (ETNs) and find that such products have reached high trading volumes (up to $5 billion per day in 2012). Our goal is to provide evidence on the success of alternative hedging strategies to manage the large risk exposures in a market that has witnessed tremendous growth in recent years.

Prices of VIX-related products are driven by the dynamics of the underlying index, but recent findings also stress the importance of other, latent, state variables such as stochastic volatility-of-volatility (vol-of-vol), a stochastic mean-reversion level or jumps (see Kaeck & Alexander, 2013; Mencía & Sentana, 2013).[3] In this paper, we study the performance of dynamic hedging strategies that account for such risk factors. Our focus on hedging performance is driven by a range of considerations, including evidence in the literature that such tests are more powerful than standard option pricing exercises (Bates, 2003; Branger, Krautheim, Schlag, & Seeger, 2012; Dennis & Mayhew, 2009). Previous research on equity indices finds that sophisticated option pricing models often dominate simpler models in pricing exercises, but may lead to inferior hedging performance.[4] Building on these earlier findings our goal is to uncover the extent to which recently identified VIX option risk factors are important to improve hedging performance.

* Corresponding author.
E-mail addresses: ak486@sussex.ac.uk (A. Kaeck), n.j.seeger@vu.nl (N.J. Seeger).

[1] For figures on trading volume of VIX futures and VIX options see www.cboe.com.

[2] See www.cboe.com

[3] Earlier work focusing on simpler model specifications include Whaley (1993), Grunbichler and Longstaff (1996), Detemple and Osakwe (2000) or Dotsis, Psychoyios, and Skiadopoulos (2007).

[4] See for instance, Bakshi, Cao, and Chen (1997).

https://doi.org/10.1016/j.ejor.2019.11.034
0377-2217/

Following the significant growth in the VIX derivative market, the study of the data-generating process of the VIX index, in addition to its impact on pricing and risk management tasks is instrumental for the understanding of this still relatively new market. Now, market participants not only have access to a range of standard derivatives but also to trading strategies that offer exposure to various stylized facts; for example, the high roll yield that results from the typically upward sloping term structure of VIX futures. Driven by the increased liquidity in this market, the number of academic studies focusing on the VIX index and its futures and options has also increased significantly. Mencía and Sentana (2013) study the pricing of VIX derivatives during the sample period from 2006 to 2010 and conclude that the most successful pricing model for VIX derivatives is highly complex, including both jump components, as well as stochastic volatility-of-volatility and a stochastic long-term mean-reversion level. Bardgett, Gourier, and Leippold (2019) study the consistency of S&P 500 and VIX derivatives in a multi-factor volatility model extension of Egloff, Leippold, and Wu (2010). Park (2016) extends pricing models in Mencía and Sentana (2013) to asymmetric volatility.

Derivative markets provide a rich data source to study the dynamics of the underlying asset, as they provide information about the dynamics under both the real-world (physical) measure, as well as under some risk-neutral pricing measure (see Broadie, Chernov, & Johannes, 2007; Eraker, 2004). Option prices are also highly informative about the evolution of state variables such as stochastic volatility or jumps. Therefore, a large body of research focuses on option pricing tests to distinguish between alternative model dynamics (see for instance, Andreou, Charalambous, & Martzoukos, 2008; Bandi & Bertsimas, 2014; Christoffersen, Heston, & Jacobs, 2009; Christoffersen, Jacobs, & Mimouni, 2010; Fu et al., 2016; Wong & Lo, 2009 and others).[5] Dennis and Mayhew (2009), however, point out that discrete prices and other microstructural frictions may lead to substantial biases in standard option pricing tests. Observation errors significantly reduce the power of standard tests, even in liquid markets. They suggest that *"a productive direction for future research would be to focus on directly testing option models by testing the accuracy of the assumed equation describing the dynamics of the underlying asset, and, in models where it would be appropriate to do so, testing the performance of dynamic replicating strategies."*[6] Similarly, Bates (2003) points out that because of the high serial correlation in implied volatility surfaces, out-of-sample hedging tests are preferred over short-term out-of-sample pricing tests. The idea of comparing alternative option pricing models using empirical hedging errors is not new to the literature (see Bakshi et al., 1997; Dumas, Fleming, & Whaley, 1998; Fabozzi, Paletta, Stanescu, & Tunaru, 2016). Nevertheless the most frequently applied yardstick for measuring model performance in the empirical option literature to date remains the study of option pricing errors.

Empirical hedging exercises are out-of-sample tests. Hedging portfolios are formed using market information on day t, as well as model-based information such as structural model parameters, hedge ratios or estimates of latent state variables. The performance of the hedge is evaluated at time $t+h$ when the hedge is lifted or rebalanced. This requires no additional model-based information and the hedging error obtained from such exercises is not biased towards models with a larger number of parameters or state variables. Pricing exercises, on the other hand, typically require the calibration of latent state variables on out-of-sample dates when the market price is compared to the model price. Such an approach often relies on further calibration as the model-based price depends crucially on the level of the state variables at time $t+h$.

Focusing on dynamic replication strategies has another important advantage, because such tests more closely resemble the way models are applied in practice. Prices in liquid option markets are driven by supply and demand, therefore, the main use of a model in a liquid market is to gauge the likelihood of future price movements and to set up adequate hedging strategies.[7] Following this line of research, Bakshi et al. (1997) conduct tests of alternative equity index models to investigate the impact of model dynamics on the hedging performance. Interestingly, they find that simpler models outperform in their hedging exercise, whereas more complex jump models outperform in their option pricing exercise. Branger et al. (2012) study how omitting factors affect the hedging performance. They conclude that hedging errors *"provide a useful economic measure for how well a model captures the true data-generating process."* Alexander and Nogueira (2007) point out that taking into account the leverage effect is crucial for equity index hedge performance. Interestingly, the simple Black–Scholes (BS, Black & Scholes, 1973) model is a serious competitor in equity markets, and various adjustments of simple Black–Scholes hedging strategies have been proposed (see Derman, 1999).

To address our research questions, we first propose a new VIX futures and option pricing model by extending the dynamics in Mencía and Sentana (2013) to correlated state variables. Considering the importance of the leverage parameter for the performance of equity index hedges, we explicitly allow for a correlation between the VIX index and its stochastic variance, as well as a correlation between the index and its long-term mean-reversion level.[8] Equipped with a very general modeling framework we derive minimum-variance hedge ratios for these models. Such hedge ratios are required as modeling stochastic variance and jumps leave the market incomplete and, therefore, hedge ratios that perfectly hedge the exposure to the underlying are not available. As a solution for this situation, Bakshi et al. (1997) derive minimum-variance hedge ratios for equity options in the context of stochastic variance and jump models which take into account the correlation in the state-variables as well as the impact of jumps on the dynamic hedging strategy. We demonstrate that the existence of jumps leads to a substantial adjustment of hedge ratios as jumps in the VIX index are typically relatively large.

In our empirical study we employ a data set of VIX futures and options from January 2007 to December 2014.[9] For our main analysis we estimate model parameters in-sample using 2007 and 2008 data use the remaining sample (2009–2014) for out-of-sample pricing and hedging exercises. In robustness tests we also run pricing and hedging performance tests for different subperiods and for frequently recalibrated models. Our empirical findings can be summarized as follow. First, we find only marginal differences in the hedging performance of alternative diffusive option pricing models that account for stochastic vol-of-vol and a stochastic long-term mean-reversion level. This result differs from recent pricing evidence in Park (2016) who shows that accounting for asymmetric

[5] Model uncertainty for variance swaps and forward start options in a stochastic volatility and jump models has been studied in Coqueret and Tavin (2016).

[6] Dennis and Mayhew (2009), p.190.

[7] Models become important to price products in illiquid markets although such prices are rarely publicly available for research.

[8] In a contemporaneous paper, Park (2016) proposes a similar model, which is however restricted to only a correlation between the VIX and its variance process. He finds strong support for such feature using option pricing metrics. Our paper differs from Park (2016) as we present limited pricing evidence and instead focus on the hedging performance.

[9] Since February 2006 is the first month of VIX options being traded we opt for using data from 2007 onward to make sure that there is enough liquidity in the VIX option market to see the data affected by the smallest amount of market microstructure noise possible. VIX ETFs could also be employed as hedging instruments, but as the maturity of VIX futures and options coincides, we do not explore this further in our paper.

stochastic variance of the VIX leads to significant pricing improvements; a result that we confirm in our own pricing errors study. Adding the possibility of jumps leads to no further improvements and our findings support that simple models are not outperformed by more advanced specifications. Interestingly, we show that the best performing hedge is given by the Black model (using the implied volatility of each option to calculate hedge ratios). This simple model significantly outperforms all structural models in our out-of-sample period from 2009 to 2014. The out-performance of the Black hedge is highly robust over time, and across hedging horizons. Given this strong finding, we explore several different avenues. First, we test how robust our results are with respect to the estimation methodology but find that the Black model remains the best-performing strategy independent of how structural models are estimated. We also explore the possibility of regime-specific behavior and apply the volatility regimes of Derman (1999) to the VIX option market. We find that the Black model still provides the most successful and consistent hedging strategy. Finally, we discuss possible explanations for our results, which include overfitting of sophisticated option pricing models due to noisy price observations. In additional empirical tests, we show that the Black hedge is highly robust in the presence of noisy market data.

The remainder of the paper is structured as follows. Section 2 presents alternative VIX option pricing models and derives option prices as well as hedge ratios. Section 3 introduces the data and the model estimation methodology. We present our empirical results in Section 4. Section 5 concludes.

2. VIX derivative pricing models and hedge ratios

2.1. Benchmark model

We employ an extension of previously proposed option pricing models as our benchmark. First, we model the volatility of the VIX index as a stochastic process as there is strong evidence to suggest that this feature is important for both the time series of the VIX index as well as VIX derivatives (see Kaeck & Alexander, 2013; Mencía & Sentana, 2013). Mencía and Sentana (2013) also find empirical support for a stochastic long-term mean-reversion level for the VIX index. Recently, Park (2016) builds on these findings and highlights the importance of a non-zero correlation coefficient between VIX index changes and their variance, as well as the importance of modeling upward jumps in VIX index values.

Our benchmark specification extends models in the literature by allowing the long-term mean-reversion level to be correlated to the VIX index. To remain within the affine model class of Duffie, Pan, and Singleton (2000), we assume that under the risk-neutral measure $\mathbb{Q}$ the (natural) logarithm of the VIX index $v_t = \ln(\mathrm{VIX}_t)$ follows[10]

$$dv_t = \kappa_v(m_t - v_t)dt + \sqrt{x_t}\,dW_t^v + \rho_m dW_t^m + \xi_t dN_t \tag{1}$$

$$dx_t = \kappa_x(\theta_x - x_t)dt + \sigma_x\sqrt{x_t}\left(\rho_x dW_t^v + \sqrt{1-\rho_x^2}\,dW_t^x\right) \tag{2}$$

$$dm_t = \kappa_m(\theta_m - m_t)dt + \sigma_m dW_t^m, \tag{3}$$

where x_t drives the diffusive variance of v_t and m_t denotes the stochastically varying mean-reversion level. The parameters ρ_x and ρ_m drive the covariation between v_t and the latent state variables. The parameters κ_v, κ_x and κ_m are the speed of mean-reversion to the long-term values m_t, θ_x and θ_m, respectively. The parameters σ_x and σ_m drive the size of the diffusion term of the two latent state variables. We allow for the possibility of jumps in v_t and follow a simple structure where ξ_t is normally distributed with mean μ_J and standard deviation σ_J. The probability of a jump in the Poisson process N_t is given by $\mathbb{Q}(dN_t = 1|\mathcal{F}_t) = \lambda_v v_t dt = \lambda_t dt$, where λ_v is a positive constant.[11] For ease of notation we collect all structural model parameters in the vector Θ.

Since the proposed model is affine in the state variables v_t, m_t and x_t, the characteristic function $f^v_{t,T}(u) = E_t^{\mathbb{Q}}[\exp(u\,v_T)]$ for $u \in \mathcal{D} \subset \mathbb{C}$, where $\mathcal{D}$ is the set for which the expectation is well-defined, is given by a log-linear form (see Duffie et al., 2000):[12]

$$\ln f^v_{t,T}(u) = \mathcal{A}(\tau,u,\Theta) + \mathcal{B}(\tau,u,\Theta)v_t + \mathcal{C}(\tau,u,\Theta)x_t + \mathcal{D}(\tau,u,\Theta)m_t$$

where $\tau = T - t$. Throughout the remainder we drop the explicit dependence on Θ and, when appropriate, use shorthand notation (for fixed parameters and fixed $u \in \mathbb{C}$), for instance $\mathcal{A}(\tau) \equiv \mathcal{A}(\tau,u,\Theta)$. Jump sizes ξ_t are independent over time and are defined in terms of their jump transform $f^{\xi}(\nu) = E^{\mathbb{Q}}[\exp(\nu\xi_t)]$ for $\nu \in \mathbb{C}$. To derive prices of VIX derivatives in our benchmark model, we require an analytic expression for the characteristic function. A straightforward application of Proposition 1 in Duffie et al. (2000) provides a set of (complex-valued) ordinary differential equations (ODEs) as follows:

Proposition 1. *Given the dynamics in Eqs. (1)–(3), the complex-valued functions $\mathcal{A}(\tau) \equiv \mathcal{A}(\tau,u,\Theta)$, $\mathcal{B}(\tau) \equiv \mathcal{B}(\tau,u,\Theta)$, $\mathcal{C}(\tau) \equiv \mathcal{C}(\tau,u,\Theta)$ and $\mathcal{D}(\tau) \equiv \mathcal{D}(\tau,u,\Theta)$ satisfy the following set of complex-valued ODEs:*

$$\begin{aligned}
\frac{d\mathcal{A}(\tau)}{d\tau} &= \theta_m\kappa_m\mathcal{D}(\tau) + \theta_x\kappa_x\mathcal{C}(\tau) + \sigma_m\rho_m\mathcal{B}(\tau)\mathcal{D}(\tau) \\
&\quad + \frac{1}{2}\sigma_m^2\mathcal{D}(\tau)^2 + \frac{1}{2}\rho_m^2\mathcal{B}(\tau)^2 \\
\frac{d\mathcal{B}(\tau)}{d\tau} &= -\kappa_v\mathcal{B}(\tau) + \lambda_v\left(f^{\xi}(\mathcal{B}(\tau)) - 1\right) \\
\frac{d\mathcal{C}(\tau)}{d\tau} &= -\kappa_x\mathcal{C}(\tau) + \frac{1}{2}\mathcal{B}(\tau)^2 + \frac{1}{2}\sigma_x^2\mathcal{C}(\tau)^2 + \rho_x\sigma_x\mathcal{B}(\tau)\mathcal{C}(\tau) \\
\frac{d\mathcal{D}(\tau)}{d\tau} &= \kappa_v\mathcal{B}(\tau) - \kappa_m\mathcal{D}(\tau)
\end{aligned}$$

subject to the boundary conditions $\mathcal{A}(0) = \mathcal{C}(0) = \mathcal{D}(0) = 0$ and $\mathcal{B}(0) = u$.

The ODE for $\mathcal{B}$ can be solved analytically if $\lambda_v = 0$, the remaining ODEs need to be solved numerically. As in Park (2016), we use standard numerical ODE solvers.

To obtain quasi-analytical expressions for the prices of VIX derivatives, we use Fourier transform methods similar to Ballotta, Deelstra, and Rayée (2017).[13] Following Bates (2006), VIX futures and options in our models can be obtained as follows:[14]

[10] Alternatively we could model the dynamics of the VIX index rather than its logarithm. We build on existing evidence in Kaeck and Alexander (2013), Bao, Li, and Gong (2012) and Mencía and Sentana (2013) that modeling the log rather than the index itself is empirically more successful.

[11] There are many alternative (and equally tractable) modeling approaches for jumps, such as extensions to infinite-activity Levy-models such as in Carr, Geman, Madan, and Yor (2002), Corsaro, Kyriakou, Marazzina, and Marino (2019), Fusai, Germano, and Marazzina (2016), and Mercuri and Rroji (2018) or other jump size distributions such as in Kou (2002). We focus on a simple and intuitive model here and comment on extensions further below.

[12] We follow standard notation where $E_t^{\mathbb{Q}}[X]$ denotes the time-t conditional expectation of a random variable X under the measure $\mathbb{Q}$. Furthermore, $\mathbb{C}$ is the set of complex numbers.

[13] Since we model the log of the VIX index, our pricing problem differs only marginally from pricing European options on equity indices for which state-of-the-art pricing models are also affine in the logarithm of the underlying index. European options have the advantage of quasi closed-form option pricing formulae, in contrast to the numerical procedures that need to be applied to American options, see Fabozzi, Paletta, and Tunaru (2017), Chockalingam and Muthuraman (2015), and Jin, Li, Tan, and Wu (2013) for recent advances.

[14] The proof is provided in the appendix.

Proposition 2. *Given the dynamics in Eqs. (1)–(3), the price of a VIX futures contract with maturity τ is given by*

$$F(v_t, x_t, m_t, \tau) = e^{\mathcal{A}(\tau,1)+\mathcal{B}(\tau,1)v_t+\mathcal{C}(\tau,1)x_t+\mathcal{D}(\tau,1)m_t}.$$

The price of a European call option on the VIX index with strike K and time to maturity τ is given by

$$C(v_t, x_t, m_t, K, \tau) = e^{-r\tau}\left(F(v_t, x_t, m_t, \tau) - \frac{1}{2}K - \frac{K}{\pi}\int_0^\infty \Re[g(o, v_t, x_t, m_t, \tau, K)]do\right).$$

where $g : \mathbb{R}^6 \mapsto \mathbb{C}$ is defined as

$$g(o, v_t, x_t, m_t, \tau, K) = \frac{e^{-io\ln(K)+\mathcal{A}(\tau,io)+\mathcal{B}(\tau,io)v_t+\mathcal{C}(\tau,io)x_t+\mathcal{D}(\tau,io)m_t}}{io \times (1-io)}$$

and $\Re$ denotes the real part of a complex number.

Note that the formula given here has the advantage over the pricing formula used by Park (2016) in that it requires only one numerical integral for the price of a call option. Put option prices follow from a simple application of put-call parity.

Our general model nests a range of simpler specifications, and this allows us to identify the empirical relevance of several model features. In this paper, we focus on six different specifications, depending on (a) whether jumps in the VIX index are modeled, and (b) the assumptions on the correlation between state variables. The full model with normally distributed jumps is labeled SVVJC2, the model with zero-correlation between state variables (i.e., imposing the restriction $\rho_x = \rho_m = 0$) is labeled SVVJ, whereas the specification with only a non-zero correlation between v_t and x_t is labeled SVVJC1 (i.e. imposing only the restriction $\rho_m = 0$).[15] The same three model classes are tested for models without jumps (i.e., imposing $\lambda_v = 0$). These models are labeled SVV, SVVC1 and SVVC2.

2.2. Hedge ratios and hedging errors

Our reduced-form specification in Eqs. (1)–(3) implies an incomplete market model and, hence, perfect hedging is not possible (even in continuous time). We employ a standard (local) risk minimization strategy and use VIX futures contracts as a hedging instrument to minimize the variance of the hedging error (see for instance, Alexander & Nogueira, 2007 or Bakshi et al., 1997). This approach has two major benefits. First, it adjusts standard (partial derivative) hedge ratios by taking into account the correlation between state variables and, hence, a partial hedge of the risk arising from risk factors that are correlated with the hedging instrument can be achieved. And second, as jump risk contributes to the variance of the hedging portfolio, the minimum-variance hedge ratios also adjust for jump risk. We summarize our main result in the next proposition.[16]

Proposition 3. *In the diffusion benchmark model, the local minimum-variance hedge ratio for a VIX call option with strike K and time-to-maturity τ (at time t) is given by*

$$\Delta_t^c(v_t, x_t, m_t, K, \tau) = \left[\left(\mathbf{F^d}\right)'\mathbf{VF^d}\right]^{-1}\left(\mathbf{F^d}\right)'\mathbf{VC^d} \quad (4)$$

where $\mathbf{F^d} = [F_v, F_x, F_m]'$, $\mathbf{C^d} = [C_v, C_x, C_m]'$ and subscripts denote partial derivatives of the VIX futures and call option price formula of Proposition 2, respectively. **V** *is the matrix of quadratic covariations defined as*

$$\mathbf{V} = \frac{1}{dt} \times \begin{bmatrix} d\langle v_t, v_t\rangle & d\langle v_t, x_t\rangle & d\langle v_t, m_t\rangle \\ d\langle x_t, v_t\rangle & d\langle x_t, x_t\rangle & d\langle x_t, m_t\rangle \\ d\langle m_t, v_t\rangle & d\langle m_t, x_t\rangle & d\langle m_t, m_t\rangle \end{bmatrix}.$$

In models with jumps in the VIX index the local minimum-variance hedge ratio is given by

$$\Delta_t^c(v_t, x_t, m_t, K, \tau) = \left[\left(\mathbf{F^d}\right)'\mathbf{VF^d} + A\right]^{-1}\left[\left(\mathbf{F^d}\right)'\mathbf{VC^d} + B\right] \quad (5)$$

where

$$A = \lambda_t \times F(v_t)^2\left(f^\xi(2\mathcal{B}(\tau,1)) - 2f^\xi(\mathcal{B}(\tau,1)) + 1\right)$$
$$B = \lambda_t \times \left\{\overline{FC} - F(v_t)\bar{C}_1 - C(v_t)\bar{F}_1 + C(v_t)F(v_t)\right\}$$

with

$$\bar{F}_1 = F(v_t)f^\xi(\mathcal{B}(\tau,1)), \qquad \bar{F}_2 = F(v_t)^2 f^\xi(2\mathcal{B}(\tau,1))$$
$$\bar{C}_1 = e^{-r\tau}\left(\bar{F}_1 - \frac{1}{2}K - \frac{K}{\pi}\int_0^\infty \Re\left[f^\xi(\mathcal{B}(\tau,io)) \times g(o,\tau,v_t,x_t,m_t)\right]do\right)$$
$$\overline{FC} = e^{-r\tau}\left(\bar{F}_2 - \frac{1}{2}K\bar{F}_1 - \frac{KF(v_t)}{\pi}\int_0^\infty \Re\left[f^\xi(\mathcal{B}(\tau,io) + \mathcal{B}(\tau,1)) \times g(o,\tau,v_t,x_t,m_t)\right]do\right).$$

The time-t minimum-variance hedge ratio of a put option with strike K and maturity τ is given by:

$$\Delta_t^p(v_t, x_t, m_t, K, \tau) = \left[\left(\mathbf{F^d}\right)'\mathbf{VF^d} + A\right]^{-1}\left[\left(\mathbf{F^d}\right)'\mathbf{VC^d} + B - e^{-r\tau}\left[\left(\mathbf{F^d}\right)'\mathbf{VF^d} + A\right]\right].$$

Proposition 3 allows us to calculate hedge ratios for all models considered in this paper. Comparing our results to the literature on hedging equity indices (see for instance, (Alexander & Nogueira, 2007)), we highlight some interesting differences. For equity indices the minimum-variance delta requires two adjustments to the standard (partial derivative) delta. First, any latent state variable that is correlated to the index leads to an adjustment term as part of the state variable risk can be hedged with the underlying. And second, there is an adjustment for the possibility of jumps. Proposition 3 implies that in our setup, there is a further adjustment to the standard delta that arises due to the explicit dependence of the hedging instrument (the VIX futures contract) on the latent state variables x_t and m_t.[17] We summarize the difference between minimum-variance hedge ratios and partial derivative hedge ratios in the next proposition.

Proposition 4. *In the diffusion models, if VIX futures prices satisfy $\mathcal{C}(\tau,1) = \mathcal{D}(\tau,1) = 0$ for all τ, and the latent state variables x_t and m_t are uncorrelated to v_t, then the minimum-variance delta and the standard (partial derivative) delta hedge ratio are the same.*

In Fig. 1 we provide an example of the difference between standard-delta and minimum-variance-delta hedge ratios for VIX call options in two models (SVVC2 and SVVJC2). These plots are calculated for strike levels from 15 to 40 with $v_t = \ln(25)$, $x_t = 1$ and $m_t = \ln(20)$. The call option maturity is two weeks, and the structural parameters are from Table 2 below. The risk-free rate is assumed to be 3%. In the SVVC2 model, the two hedge ratios

[15] As we expect the correlation between the stochastic variance x and v_t to be of first-order importance, we restrict our empirical study to this model.

[16] The proof is provided in the appendix.

[17] Note that this is not the case for equity indices where the futures contract formula is not a function of stochastic volatility or other state variables.

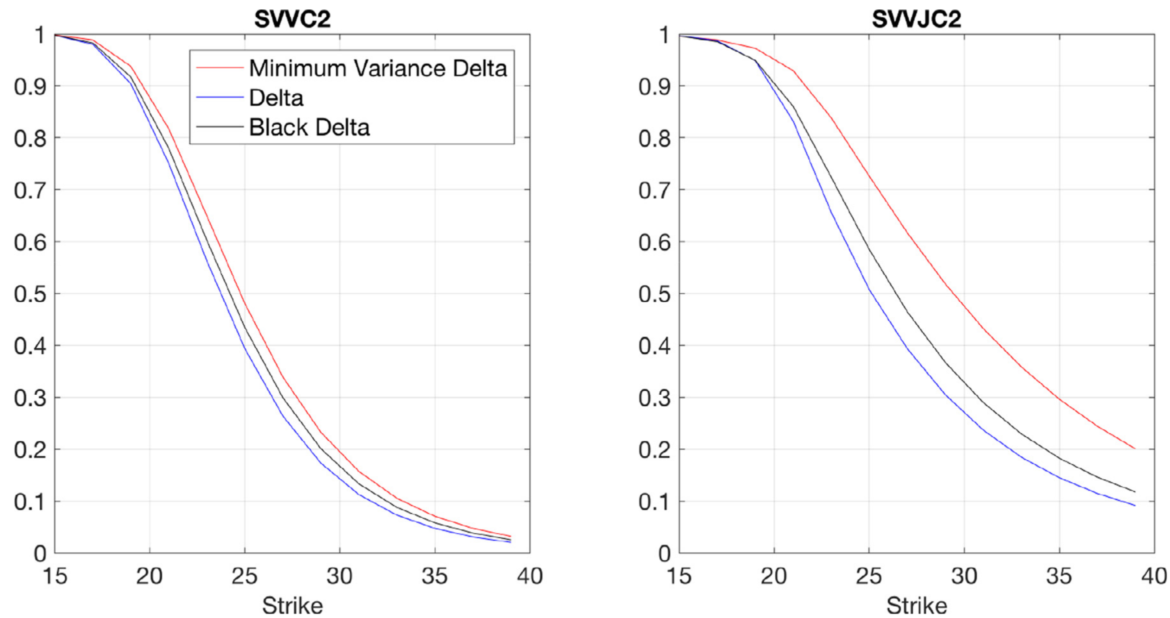

Fig. 1. Hedge Ratios. This figure shows minimum-variance and standard delta hedge ratios for two models (SVVC2 and SVVJC2) for strike levels from 15 to 40 for $v_t = \ln(25)$, $x_t = 1$ and $m_t = \ln(20)$. The hedge ratios are for option with a maturity of two weeks, assuming the parameters are taken from Table 2 below. The risk-free rate is assumed to be 3%. Black model deltas are provided in black and use the implied volatility of the option for hedge ratio calculations.

show substantial differences, especially for at-the-money options. For a strike of 25, for instance, the difference between the partial delta and the local-risk minimization delta is roughly 10%, whereas hedge ratios are similar for far in-the-money and far out-of-the-money option contracts. In the right graph of Fig. 1 we see that the difference between the partial derivative and the minimum-variance delta becomes more substantial for models with an additional jump component.

2.3. Measuring the hedging performance

To measure the empirical performance of alternative models we build hedging portfolios as follows. On each trading day t, and for each option in the sample, we denote the model-based hedging error for a call option i at time t over a time interval h (i.e. from time t to time $t+h$) as

$$\mathrm{HE}_{i,t} = C^M(K, \tau - h, t+h) - C^M(K, \tau, t) - \Delta_t^c(v_t, x_t, m_t, K, \tau) \times (F^M(t+h, \tau - h) - F^M(t, \tau)), \tag{6}$$

where C^M and F^M denote market prices of VIX call options and VIX futures, respectively. Hedge errors for put options are defined equivalently, with an obvious adjustment in the notation of Eq. (6). We use standard intervals of one day and one week in the empirical analysis that follows. Given the empirical hedge errors over the whole sample period, we then calculate the root-mean-square hedging error (RMSHE), as our goal is to identify the model with the smallest hedge error risk. Where the RMSHE is defined as

$$\mathrm{RMSHE}_t = \sqrt{\frac{1}{N_t^o} \sum_{i=1}^{N_t^o} \mathrm{HE}_{i,t}^2}, \tag{7}$$

with N_t^o denoting the number of VIX options in the sample on day t. Other hedging error definitions, such as relative hedging errors that normalize $\mathrm{HE}_{i,t}$ by the current market price, have been used in the empirical option pricing and hedging literature (see Christoffersen & Jacobs, 2004; Date & Islyaev, 2015). We will explore such alternatives below as a robustness check.

2.4. Black model

For comparison, and because it is the industry benchmark, we also investigate the performance of the standard Black model (Black, 1976). The model dependent hedge ratio in Eq. (6) is replaced by the Black delta

$$\Delta_t^{c,b}(\sigma, F^M(t,\tau), K, \tau) = e^{-r\tau} N(d_1) \quad \text{with}$$
$$d_1 = \frac{\ln(F^M(t,\tau)/K) + (\sigma^2/2)\tau}{\sigma\sqrt{\tau}}, \tag{8}$$

where σ denotes the volatility of the underlying and $F^M(t, \tau)$ denotes the time-t VIX futures market price with time to maturity τ and strike price K.[18] The Black delta can be calculated from market observables only and does not require any model calibration. We follow standard practice and use the implied volatility for each option in the calculation of the Black delta $\Delta_t^{c,b}(IV^M(K, F^M(t,\tau), \tau), F^M(t,\tau), K, \tau)$ where $IV^M(K, F^M(t, \tau), \tau)$ denotes the time-t market implied volatility of a call option with strike K and time to maturity τ. This use of the Black model hedge ratios is inconsistent with a unique risk-neutral pricing measure as all options have their unique implied volatility. Schönbucher (1999) discusses the theoretical restrictions that ensure the absence of arbitrage in related market models of implied volatility. The Black delta is an established benchmark in the option hedging literature (see also Hull & White, 2017), due to its overall empirical success, its simplicity and its ubiquitous use in the financial industry. Our calculation of the Black delta is in line with many recent studies in the empirical option hedging literature, earlier studies such as Bakshi et al. (1997) use a single volatility parameter that is recalibrated daily, independent of the option's maturity and strike.

3. Data and methodology

3.1. Data

We collect VIX option data from the Chicago Board Options Exchange (CBOE). VIX options started trading in February 2006 and

[18] Put deltas follow from put-call parity.

our sample ends in December 2014. Since VIX options have been more illiquid during the first months of trading, we start our data sample in January 2007. The data cleaning steps described in the following are needed to drop sparsely traded option observations from the data set since those might represent stale prices that do not reflect most recent economic information.[19]

To ensure that our results are based on reliable quotes we first remove quotes for which the mid-price violates standard no-arbitrage conditions, such as in Bakshi et al. (1997).[20] We then discard quotes for extreme moneyness levels as these are only infrequently traded and may have stale prices. We define moneyness as K/F (for further details on trading volumes and other statistics of the data, see below). The moneyness definition implies that for $K > F$ we refer to in-the-money (ITM) put and out-of-the-money (OTM) call options, for $K < F$ we refer to out-of-the-money put and in-the-money call options, and for $K \approx F$ we refer to at-the-money (ATM) put and call options. After an initial screening of the trading volumes of different option categories, we opt to include options with moneyness levels ranging from 60% to 160%.[21] Finally, we remove option quotes for which the database records a trading volume of zero.

To ensure that our results are not based on obvious recording errors we apply an additional filter to the data. For each day in the sample and each time-to-maturity we first fit a cubic polynomial to all market implied volatilities that have survived the first filtering stage. With this flexible parametric form, we expect individual data points not to deviate too strongly from the fitted curve. Subsequently, we discard all quotes which deviate either more than three error term standard deviations or ten implied volatility points from the fitted curve. After removing such quotes we re-fit the polynomial to check whether further quotes should be discarded and continue this procedure until all quotes are within the specified boundaries. The boundaries are very wide and the procedure only rarely indicates that a quote is to be removed. The aim of the additional filter is merely to ensure that we do not rely on quotes that may very likely be recording errors. This is important as we do not want the hedging performance to be dominated by implausible outliers in the data.

Table 1 provides summary statistics for our raw option data. First, this table justifies the data filters presented earlier as contracts not included in the final dataset have either substantially lower trading volume or are excluded on other grounds. Overall, the implied volatilities show a smile-shape pattern known from many other markets. Implied volatilities tend to be lower for options close to ATM, and increase the more the strike deviates from the current futures price. In addition, the skew of the implied volatility curve is more extreme for short maturities, where implied volatility is also on average higher than for options with longer-time-to maturity. This is expected as long-term futures have substantially less uncertainty about their average volatility over the life of the contract than short-term futures (Alexander & Korovilas, 2013). The trading volume of call and put options is roughly symmetric for ATM options, and OTM options are more liquid than ITM options.

VIX futures for the same sample period are collected from the CBOE website. For each trading day we obtain all maturities that are quoted. To filter out unreliable quotes, we follow a similar procedure as for the options.[22] Our risk-free rate is collected from OptionMetrics. OptionMetrics provides readily employable zero-bond interest rates for various time horizons (zero-curve) that are calculated from ICE IBA LIBOR rates and settlement prices of CME Eurodollar futures. Using the zero curve from OptionMetrics allows simple matching of options with the respective risk-free interest rates based on an option's time-to-maturity.

3.2. Parameter estimation and state variable filtering

The main challenge in estimating option pricing models stems from the fact that option prices depend on state variables, such as stochastic volatility in a non-linear way. For this reason, standard filtering techniques such as the (standard) Kalman filter are not applicable. Mencía and Sentana (2013) use the extended Kalman filter to estimate VIX option pricing model parameters and to filter out unobserved state variables, whereas Park (2016), following Carr and Wu (2007), employ the unscented Kalman filter of Julier and Uhlmann (1997) and Wan and Van Der Merwe (2000). We follow the same approach in this paper as the unscented Kalman filter has been shown to provide an accurate filtering and estimation tool in the context of non-linear state-space models (see Christoffersen, Dorion, Jacobs, & Karoui, 2014).[23] We also provide empirical results using alternative model estimation procedures and present these in the robustness section.

To cast our problem in state-space form we assume that both VIX option prices as well as futures prices are observed with error. Our observation equations on day t are therefore given by

$$IV^M(K_i, F^M(t,\tau_i), \tau_i, \omega_i) = IV(\nu_t, x_t, m_t, K_i, \tau_i, \omega_i) + \varepsilon^o_{i,t} \qquad i = 1, \dots, N^o_t$$

$$\ln F^M(t,\tau_i) = \ln F(\nu_t, x_t, m_t, \tau_i) + \varepsilon^f_{i,t} \qquad i = 1, \dots, N^f_t$$

where $IV^M(K_i, F_{t,\tau_i}, \tau_i, \omega_i)$ denotes the market-implied Black volatility of option i on day t, $IV(\nu_t, x_t, m_t, K_i, \tau_i, \omega_i)$ is the model-implied Black volatility of option i on day t, and $F(\nu_t, x_t, m_t, \tau_i)$ is the model price of a futures contract at time t with maturity τ_i. We use ω_i to indicate whether the option is a call or a put option. We assume iid, zero-mean error terms $\varepsilon^o_{i,t}$ and $\varepsilon^f_{i,t}$ with standard deviations σ_o and σ_f. The number of VIX options and VIX futures in the sample on day t are given by N^o_t and N^f_t, respectively.

To estimate the models we require assumptions on the dynamics of the latent state variables under the real-world measure $\mathbb{P}$. We assume that the SDEs under $\mathbb{P}$ are given as follows:

$$dx_t = \kappa^P_x\left(\theta^P_x - x_t\right)dt + \sigma_x\sqrt{x_t}dW^{\nu,P}_t \quad \text{and}$$
$$dm_t = \kappa^P_m\left(\theta^P_m - m_t\right)dt + \sigma_m dW^{m,P}_t,$$

[19] See Bakshi et al. (1997), Mayhew (2002), Bondarenko (2014), Jackwerth and Rubinstein (1996), Bates (2000), and (Christoffersen, Jacobs, & Ornthanalai, 2012)). The effect of sparsely traded options can even be amplified by market microstructure effects such as large bid-ask spreads (see (Mayhew, 2002)). Large bid-ask spreads are set by market makers to account for illiquidly traded options that they carry in their inventory. Those options pose a risk to them since stale prices do not reveal the true fundamental value at a specific moment in time, which can lead to unfavorable trades with investors that have superior insider information. To counteract such risks, market makers pose large bid-ask spreads that cover for possible losses from unfavorable trades. Such large bid-ask spreads make it difficult to determine the true current fundamental value of an option. Our data selection does not completely free the option prices from illiquidity and market microstructure effects, but provides us with a data set which is less affected by them.

[20] The mid-price is defined as the average of the bid and ask price.

[21] In other words, we remove from our dataset very far in-the-money put and call options, and very far out-of-the-money put and call options since those options are less liquidly traded. We also remove options with very short or long time-to-maturity as these also suffer from the same liquidity issues. We keep options with a maturity greater than one week and less than half a year.

[22] We first fit a cubic polynomial to the data and remove futures quotes that deviate either more than three error term standard deviations or more than three volatility points. As before, if this procedure results in a quote identified as an outlier, we repeat the procedure until all futures quotes are within the specified boundaries.

[23] We give an overview of the unscented Kalman filter in Section A.4 of the Appendix. For full details on the unscented Kalman filter we refer the reader to Hirsa (2016).

Table 1
Summary Statistics (VIX Options). This table reports summary statistics for the VIX option sample from January 2007 until December 2014. We provide the number of contracts, the average implied volatility and the average trading volume of put and call options for various moneyness categories (shown in the first column) and different time-to-expiration categories (shown in the first row). The first line for each category provides the number of contracts in the database, overall the dataset includes 218,408 options. The second line (the number in brackets) provides the average implied Black volatility (Black, 1976) for the option category and the two numbers in the third line provide the average daily trading volume for call and put options, respectively. Moneyness is defined as $M = K/F$ where K is the strike price and F is the VIX futures price of a contract with the same time-to-maturity as the VIX option.

Moneyness M	Days to Expiration				
	All	1–7	8–30	31–180	181–360
All	218408	6965	35152	171310	4981
	(0.7966)	(1.4773)	(1.0270)	(0.7298)	(0.5138)
	[2563, 2362]	[8567, 5850]	[6564, 4391]	[1778, 1619]	[316, 530]
$0.4 \leq M < 0.6$	3888	227	770	2763	128
	(0.9210)	(2.9473)	(1.3274)	(0.6629)	(0.4520)
	[177, 420]	[226, 0]	[410, 1392]	[111, 408]	[36, 27]
$0.6 \leq M < 0.9$	60218	1237	7409	49691	1881
	(0.6219)	(1.3062)	(0.7818)	(0.5876)	(0.4503)
	[521, 2638]	[1884, 15,752]	[1225, 8818]	[351, 2070]	[83, 667]
$0.9 \leq M < 1.1$	55688	2436	9794	42264	1194
	(0.7361)	(1.0294)	(0.8800)	(0.6918)	(0.5226)
	[3525, 4075]	[13,703, 15,069]	[8491, 8658]	[2053, 2256]	[314, 802]
$1.1 \leq M < 1.4$	52903	2013	10894	39002	994
	(0.8861)	(1.5133)	(1.1165)	(0.7977)	(0.5608)
	[4251, 836]	[15,156, 2171]	[10,587, 1329]	[2892, 439]	[432, 67]
$1.4 \leq M < 1.6$	23075	589	3754	18345	387
	(0.9608)	(2.0912)	(1.2980)	(0.8633)	(0.5926)
	[2995, 193]	[7829, 235]	[7758, 348]	[2422, 92]	[523, 18]
$1.6 \leq M < 2$	22636	463	2531	19245	397
	(1.0118)	(2.6335)	(1.4357)	(0.9252)	(0.6131)
	[1886, 99]	[2035, 82]	[5995, 115]	[1637, 96]	[856, 24]

with $\kappa_x^P\theta_x^P = \kappa_x\theta_x$ and $\kappa_m^P\theta_m^P = \kappa_m\theta_m$. These risk premium restrictions are equivalent to assuming that the change of measure for the variance process x_t is given by $dW_t^{v,P} + \eta_x\sqrt{x_t}dt = dW_t^v$ and $dW_t^{m,P} + \eta_m m_t dt = dW_t^m$. As in Carr and Wu (2007) we then discretize the state variables x and m using an Euler discretization scheme, construct the likelihood function assuming normally distributed forecasting errors and maximize the likelihood function with respect to the parameter set Θ.

4. Empirical results

We separate our sample into an in-sample period during which all structural models are estimated, and an out-of-sample period during which their hedging performance is assessed. To this end, we follow a number of papers in the empirical equity option literature, such as Ornthanalai (2014), and use Wednesday options during 2007 and 2008 for model estimation.[24] Given the estimated model parameters, the unscented Kalman filter provides a filtering tool to extract latent state variables during both the in- and out-of-sample period that can then be used to build hedging portfolios as described in Section 2.2. We provide alternative estimation methods with different estimation periods in the robustness section below.

4.1. Parameter estimates and pricing performance

Table 2 provides parameter estimates for all six competing models using the unscented Kalman filter described in Section 3.2 and Section A.4. Mean reversion levels for the VIX index κ_v are estimated between 8.16 and 11.31 depending on the

[24] Since the characteristic function needs to be evaluated numerically the calibration approach is computationally very demanding. We therefore restrict our in-sample period to weekly data to alleviate the computational demands. This has become standard in the literature, see Christoffersen et al. (2010) or Ornthanalai (2014).

Table 2
In-sample Parameter Estimation. This table reports the parameter estimation results for all models defined in Eqs. (1)–(3). The estimation period is from January 2007 to December 2008. The estimation is performed using maximum likelihood method via the unscented Kalman filter. For each parameter, we report the maximum likelihood estimates and the standard errors in parenthesis. Models are estimated to weekly VIX option and VIX futures data described in Section 3.1.

Model	SVV	SVVC1	SVVC2	SVVJ	SVVJC1	SVVJC2
κ_v	10.2848	9.2024	9.5012	9.2829	11.3108	8.1609
	(0.1401)	(0.0676)	(0.0518)	(0.0742)	(0.1136)	(0.0314)
κ_x	1.8979	2.6532	1.9931	0.9734	1.9952	1.6316
	(0.0785)	(0.0598)	(0.0466)	(0.0536)	(0.0704)	(0.0467)
θ_x	2.3218	1.7970	1.2855	0.8877	1.2076	0.5049
	(0.1419)	(0.0515)	(0.0447)	(0.0754)	(0.0539)	(0.0179)
σ_x	6.0914	3.5305	3.5100	6.4369	4.5663	7.4018
	(0.0694)	(0.0220)	(0.0449)	(0.0830)	(0.0537)	(0.0927)
κ_m	0.4594	0.4334	0.4246	0.4165	0.4595	0.4222
	(0.0076)	(0.0072)	(0.0046)	(0.0065)	(0.0068)	(0.0071)
θ_m	2.8402	2.8145	2.8431	2.5258	2.5855	2.5777
	(0.0072)	(0.0057)	(0.0059)	(0.0067)	(0.0066)	(0.0076)
σ_m	0.4225	0.3676	0.3311	0.3633	0.3872	0.4111
	(0.0141)	(0.0069)	(0.0061)	(0.0056)	(0.0045)	(0.0052)
ρ_x		0.5990	0.7997		0.7045	0.8193
		(0.0042)	(0.0077)		(0.0040)	(0.0086)
ρ_m			0.4594			-0.4078
			(0.0063)			(0.0058)
λ_v				7.2229	8.4507	6.7670
				(0.0942)	(0.1861)	(0.1101)
μ_J				0.1266	0.1123	0.0751
				(0.0022)	(0.0019)	(0.0011)
σ_J				0.1268	0.0054	0.1710
				(0.0015)	(0.0057)	(0.0011)
κ_x^P	2.5821	37.0830	35.4893	25.7241	35.0441	27.2495
	(1.3208)	(0.0800)	(0.0960)	(0.7049)	(0.2619)	(0.3735)
κ_m^P	0.6769	0.4348	0.4954	0.4882	0.5321	0.4011
	(0.0209)	(0.0159)	(0.0212)	(0.0165)	(0.0089)	(0.0264)
$\sigma_F(\%)$	2.2356	2.1095	2.0767	2.1348	2.0744	2.2178
	(0.0365)	(0.0286)	(0.0308)	(0.0262)	(0.0294)	(0.0289)
$\sigma_O(\%)$	8.5655	6.1293	5.7540	6.2530	5.8677	5.5176
	(0.0357)	(0.0207)	(0.0156)	(0.0184)	(0.0205)	(0.0173)
LogL	−19370	−17568	−17223	−17570	−17293	−16889

model assumptions. This is in line with Park (2016) whose estimates (based on an extended sample from July 2006 until January 2013) are slightly smaller with values between 6.38 and 8.96. The mean reversion parameter for the variance process x_t is also relatively stable across model specifications with values between 0.97 and 2.65. The low mean-reversion speed can be attributed to the fact that these parameters are affected by risk premia. The equivalent parameter under the real-world measure $\mathbb{P}$ tends to be much higher (for instance 35.4893 in the most general diffusion model SVVC2) and hence our results imply a significant risk premium for bearing vol-of-vol risk. Our parameter estimates for κ_x^P (for all of the models except SVV) are comparable to the time-series evidence reported in Kaeck and Alexander (2013). The parameters θ_x and σ_x show a slightly larger variation across model specifications, but overall our parameters are similar to results reported in Park (2016). The mean reversion speed for the process m_t is considerably lower than for x_t. Estimates for κ_m close to 0.50 imply an economically insignificant risk premium for long-term VIX levels. The correlation between v_t and x_t is estimated between 0.60 and 0.82, values that are in line with findings in Kaeck and Alexander (2013) who estimate a correlation coefficient of 0.66. Slightly surprising, our estimates for ρ_m are not consistent across diffusion and jump models. The estimate of 0.46 for SVVC2 implies a positive relation between the long-term mean-reversion level and the VIX index, whereas our point estimate for SVVJC2 is surprisingly negative.

As in Mencía and Sentana (2013) we find evidence for jumps. The jump parameter $\lambda_v = 7.2229$ for SVVJ implies slightly more than 20 jumps per year on average. The jumps size mean and standard deviation is estimated around 13%. Note that the parameters of the jump process are under the risk-neutral measure $\mathbb{Q}$ and it is not straightforward to interpret these results as all jump process parameters may be affected by jump risk premia. As noted by Broadie et al. (2007), theory provides very few restrictions on the transition between jump parameters under the risk-neutral and real-world measure for a Poisson jump model. Nevertheless, time-series evidence points towards similar return distribution parameters, but a lower jump probability. Kaeck and Alexander (2013) estimate 2–3 jumps per year with an average size of 0.136 and a standard deviation of 0.103. While these results provide some evidence that supports a positive jump risk premium, more definite results may be obtained from a joint estimation of the VIX index and VIX derivatives.[25]

We turn to the estimates of pricing errors. First, all models provide a similar fit to VIX futures contracts with percentage pricing errors $\sigma_F(\%)$ between 2.07 and 2.24. It is not surprising that neither jumps nor stochastic volatility plays a major role in pricing VIX futures for which the level of the state variable m_t and the level of the VIX index drive the long-end and the shape of the term structure. The pricing errors for VIX options show more variation. In the pure diffusion models without jumps, the error decreases from 8.57 to 5.75 from SVV to the specification SVVC2. Models with jumps provide further improvements, for instance SVV and SVVJ have pricing errors of 8.57 and 6.25, respectively. For SVVC2 and SVVJC2, the difference in pricing errors is, however, economically small with estimated standard deviations of 5.75 and 5.52. Unsurprisingly, the SVVJC2 model, being the model with largest number of parameters, shows the best performance in terms of in-sample model fit. In general, we find that the model with jumps always outperforms their pure diffusive counterpart as highlighted by their larger log-likelihood value. Overall we confirm findings in Park (2016) that the pricing of VIX derivatives is improved by considering pricing models with correlated state variables. Our findings regarding the importance of jumps are, however, less clear-cut.

4.2. Empirical hedging results

We first focus on a daily rebalancing exercise, where h in Eq. (6) is set to one day. We use daily data from 2009 until 2014, the full six year sample includes 119,474 observations. Table 3 provides RMSHE for the whole out-of-sample period. We first focus on the hedging performance of the diffusion models. Interestingly, the overall performance of the models is stable across different specifications with a RMSHE of 0.102–0.103. Our results are somewhat surprising as the estimated correlation between v_t and x_t is relatively large, nevertheless models with different assumptions on the correlation of state variables show very little difference in their hedging performance. These results differ from well-established empirical evidence in the equity index literature, where taking into account the correlation between state variables leads to marked improvement in the hedging performance (see Alexander & Nogueira, 2007). The hedging strategies are overall successful as they reduce the RMSHE compared to an unhedged portfolio (RMSHE: 0.444) by more than 75%. Models with additional jumps are also successful with significant RMSHE reductions, but surprisingly the additional model complexity leads to a deterioration of the hedging performance compared to models without jumps. The RMSHE in jump models ranges from 0.138 to 0.214. The overall best hedging strategy is given by the simple Black model which provides the lowest hedging errors with a RMSHE of only 0.085 representing an improvement in hedging performance of over 80% compared to the unhedged model.

The results for the more granular option categories (grouped by moneyness and time-to-maturity) in Table 3 confirm these overall results, but also reveal some interesting details. The clear outperformance of the Black hedge is most visible for option contracts with moneyness between 0.7 and 1.3, regardless of the time-to-maturity horizon. However, for the more extreme moneyness categories, 0.4–0.7 and 1.3–1.6, the outperformance of the Black hedge is less pronounced. This observation is consistent with our finding in Fig. 1 that options closer to ATM have more pronounced differences in their model dependent hedge ratios. For instance, for options with more than 90 days to maturity and moneyness between 0.4 and 0.7, we find that SVV, SVVC1, SVVC2 and SVVJC1 outperform the Black hedge, even though the outperformance is not particularly strong (RMSHE between 0.077 and 0.078 compared to 0.079 for the Black model, respectively). For options with less than 90 days to maturity and moneyness between 0.4 and 0.7 we observe that only the SVVC2 model with a RMSHE of 0.063 slightly outperforms the Black hedge with a RMSHE of 0.065. We conclude that the Black model provides a robust hedge that outperforms more sophisticated models in terms of the overall RMSHE.

To facilitate statistical comparison of the hedging performance of alternative models, we estimate model confidence sets (MCS) proposed by Hansen, Lunde, and Nason (2011). MCS use pair-wise model comparisons (based on RMSHE) as a relative performance measure between two models. The pair-wise comparison of all models leads to a ranking of the models of which the top-ranked models form the MCS according to a chosen confidence level. The models in the MCS are characterized by their performance not being significantly different from each other for the chosen confidence level, and statistically different from the performance of the models not in the MCS. Typically, confidence levels of 10% and 25% are used in the literature. Note that if the MCS is comprised of a single model then this indicates that this model shows a very strong outperformance over all other models. For a detailed definition of MCS see Section A.5 and Hansen et al. (2011).

[25] And even in this case, the equity option literature provides mixed evidence, often restricting some parameters to be identical under both the risk-neutral and real-world probability measure (see Broadie et al., 2007; Eraker, 2004; Pan, 2002).

Table 3
Hedging RMSHE (daily rebalancing). This table reports the root-mean-square hedging error (RMSHE) of the hedging errors defined in Eq. (6) for a range of different hedging strategies and different option categories. Column 1 classifies the options depending on their moneyness level, where moneyeness is defined as *K*/*F*. Column 2 classifies options according to their remaining time-to maturity in days (*DTM*: days to maturity). *Num* in column 3 provides the number of options in the option category. Column 4 provides the RMSHE for an unhedged portfolio (*Unh*) for which the delta is set to zero. Column 5 provides the RMSHE for the Black model and columns 6 to 11 for the models introduced in Section 2.1. The hedging errors calculated based on the whole out-of-sample period from January 2009 until December 2014. The re-balancing frequency is daily.

Moneyness	DTM	Num	Unh	Black	SVV	SVVC1	SVVC2	SVVJ	SVVJC1	SVVJC2
All	All	119474	0.444	0.085	0.102	0.103	0.102	0.141	0.138	0.214
0.4–0.7	All	3421	0.674	0.074	0.075	0.073	0.072	0.089	0.077	0.086
0.7–0.9	All	30474	0.508	0.078	0.097	0.089	0.086	0.125	0.113	0.140
0.9–1.1	All	37712	0.430	0.093	0.119	0.117	0.116	0.168	0.163	0.226
1.1–1.3	All	24372	0.424	0.093	0.104	0.116	0.117	0.157	0.158	0.264
1.3–1.6	All	23495	0.348	0.074	0.078	0.083	0.084	0.096	0.103	0.229
All	≤ 90	76828	0.509	0.093	0.115	0.116	0.116	0.163	0.163	0.242
0.4–0.7	≤ 90	1375	0.932	0.065	0.069	0.065	0.063	0.094	0.075	0.088
0.7–0.9	≤ 90	17808	0.606	0.083	0.110	0.100	0.096	0.146	0.134	0.161
0.9–1.1	≤ 90	25151	0.479	0.101	0.134	0.132	0.131	0.193	0.192	0.254
1.1–1.3	≤ 90	17460	0.468	0.100	0.115	0.128	0.130	0.176	0.181	0.289
1.3–1.6	≤ 90	15034	0.412	0.083	0.087	0.093	0.096	0.109	0.121	0.254
All	≥ 91	42646	0.295	0.070	0.073	0.073	0.071	0.088	0.072	0.150
0.4–0.7	≥ 91	2046	0.420	0.079	0.078	0.078	0.077	0.085	0.077	0.085
0.7–0.9	≥ 91	12666	0.322	0.071	0.077	0.072	0.071	0.089	0.073	0.105
0.9–1.1	≥ 91	12561	0.309	0.073	0.079	0.077	0.076	0.098	0.077	0.154
1.1–1.3	≥ 91	6912	0.281	0.072	0.073	0.077	0.074	0.089	0.075	0.189
1.3 – 1.6	≥ 91	8461	0.190	0.056	0.057	0.060	0.058	0.066	0.059	0.177

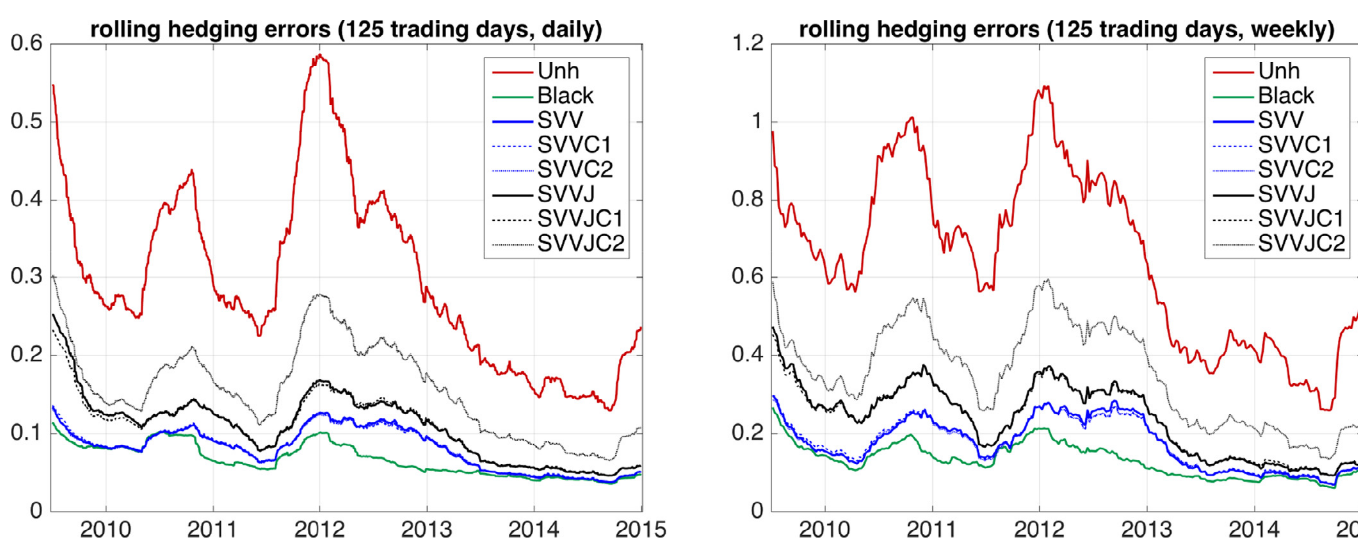

Fig. 2. Hedging Errors for Sub-Periods. This figure shows the average root-mean-square error (RMSHE) of the hedging errors defined in Eq. (6) for a range of different hedging strategies. The RMSHE are calculated as follows. Left panel: for all hedging strategies we calculate the RMSHE using all hedging errors on one trading day (aggregating over different strikes and different maturities). We then average the RMSHE over the last 125 trading days (half a year). Right panel: for all hedging strategies we calculate the RMSHE using all hedging errors on every Wednesday in the out-of-sample period (aggregating over different strikes and different maturities). We then average the RMSHE over the last 6 months and roll over on the next Wednesday. *Unh* is the hedging strategy for which the delta is set to zero, Black is the Black-delta hedge introduced in Section 2.4 and all other models are introduced in Section 2.1.

The MCS results for the daily hedging frequency, reported in Table A.1 in the appendix, confirm that the best performing hedging strategy is given by the Black delta. The Black hedge is the only model in both the 10% and 25% model confidence set, indicating strong statistical support for the Black hedge.

To investigate whether the performance difference between alternative specifications stems from a particular time period or whether the performance difference is consistent over time, we study the performance over sub-samples of our out-of-sample period from 2009 until 2014. We adopt the following approach. First, for all hedging strategies we calculate the RMSHE using all hedging errors on one trading day (aggregating over different strikes and different maturities). We then average the RMSHE over the last 125 trading days (half a year) and roll over on the next day. The left panel of Fig. 2 provides the performance of the daily hedging exercise. As is apparent from this graph, the performance of alternative model specifications is very stable over time. First, there is little difference between SVV and its extensions to correlated state variables as the hedging errors are moving in lockstep throughout the sample. The diffusion models consistently outperform jump-augmented specifications. This finding cannot be attributed to a specific time period, and is also a rather persistent observation. The Black model is the best performing model for the majority of six-month periods, merely during the beginning of 2010, average hedging RMSHE are marginally higher than for the diffusive stochastic vol-of-vol models. Overall, these results provide further evidence for the strong performance of a simple Black model hedge.

Table 4
Hedging RMSHE (weekly rebalancing). This table reports the root-mean-square error (RMSHE) of the hedging errors defined in Eq. (6) for a range of different hedging strategies and different option categories. Column 1 classifies the options depending on their moneyness level, where moneyness is defined as *K*/*F*. Column 2 classifies options according to their remaining time-to maturity in days (*DTM*: days to maturity). *Num* in column 3 provides the number of options in the option category. Column 4 provides the RMSHE for an unhedged portfolio (*Unh*) for which the delta is set to zero. Column 5 provides the hedging RMSHE for the Black model and columns 6 to 11 for the models introduced in Section 2.1. The hedging errors calculated based on the whole out-of-sample period from January 2009 until December 2014. The re-balancing frequency is weekly (Wednesday to Wednesday).

Moneyness	DTM	Num	Unh	Black	SVV	SVVC1	SVVC2	SVVJ	SVVJC1	SVVJC2
All	All	22336	0.865	0.175	0.230	0.228	0.225	0.305	0.301	0.463
0.4–0.7	All	622	1.320	0.101	0.123	0.114	0.111	0.171	0.131	0.156
0.7–0.9	All	6016	1.009	0.150	0.227	0.204	0.196	0.290	0.261	0.328
0.9–1.1	All	7119	0.841	0.197	0.270	0.262	0.259	0.358	0.354	0.493
1.1–1.3	All	4571	0.799	0.190	0.225	0.250	0.252	0.323	0.338	0.572
1.3–1.6	All	4008	0.618	0.158	0.167	0.178	0.180	0.199	0.218	0.474
All	≤ 90	13411	0.982	0.201	0.270	0.270	0.267	0.362	0.369	0.536
0.4–0.7	≤ 90	243	1.778	0.105	0.129	0.124	0.118	0.199	0.159	0.169
0.7–0.9	≤ 90	3292	1.187	0.176	0.274	0.249	0.238	0.354	0.330	0.392
0.9–1.1	≤ 90	4434	0.939	0.226	0.318	0.309	0.305	0.424	0.430	0.571
1.1–1.3	≤ 90	3074	0.876	0.210	0.251	0.284	0.287	0.366	0.395	0.629
1.3–1.6	≤ 90	2368	0.731	0.181	0.188	0.203	0.207	0.230	0.261	0.534
All	≥ 91	8925	0.649	0.126	0.151	0.143	0.142	0.189	0.147	0.324
0.4–0.7	≥ 91	379	0.911	0.099	0.119	0.108	0.107	0.151	0.110	0.146
0.7–0.9	≥ 91	2724	0.739	0.112	0.150	0.129	0.128	0.186	0.135	0.227
0.9–1.1	≥ 91	2685	0.650	0.137	0.165	0.156	0.155	0.209	0.161	0.327
1.1–1.3	≥ 91	1497	0.609	0.143	0.157	0.161	0.159	0.207	0.164	0.432
1.3–1.6	≥ 91	1640	0.403	0.118	0.130	0.134	0.130	0.144	0.135	0.370

To test whether our results depend on the rebalancing frequency, we run the same hedging study for a weekly rebalancing interval. We follow the literature and use Wednesday data. To this end, we set up a model-based hedging portfolio every Wednesday and then calculate the RMSHE when the hedge is lifted the following Wednesday. Our results for the weekly rebalancing exercise are based on 22,336 error observations for every strategy. Our main findings are summarized in Table 4, which provides RMSHEs for a variety of different option categories. Unsurprisingly, the hedging errors over one week are larger than over a daily sample, the unhedged portfolio RMSHE, for instance, increases from 0.444 to 0.865. The increase is of similar magnitude for the model-based hedges. Models do, as before, provide substantial improvements over an unhedged portfolio. Overall the model ranking carries over to the weekly hedging exercise, where the Black model dominates all other hedging models with the lowest RMSHE of 0.175. Diffusive stochastic vol-of-vol models have similar hedging performance with RMSHE ranging from 0.225 for SVVC2 to 0.230 for SVV. Models with jumps provide the worst model performance, with RMSHEs from 0.301 to 0.463. Interestingly, looking at the more detailed option categories with respect to moneyness and time-to-maturity reveals that for the weekly rebalancing the Black hedge outperforms all other models over all sub-categories. We provide model confidence sets and a plot of RMSHEs over the out-of-sample period in Table A.2 and the right panel of Fig. 2, respectively. We do not discuss these results in detail, as they provide strong support for findings discussed for a one-day rebalancing exercise. As before, we conclude that the Black model provides substantial improvements over more advanced model hedges and this result is even stronger for the weekly rebalancing case.

4.3. Alternative estimation procedures

The model estimation procedure introduced in the previous section keeps model parameters fixed during the whole sample period. This approach has the advantage of being consistent with the model assumptions of constant parameters in Eqs. (1)–(3). In addition, the unscented Kalman filter allows us to filter out state variables during the out-of-sample period by including a (log-likelihood) penalization if state variable dynamics are inconsistent with the assumed model dynamics. While this approach is theoretically sound, one may argue that the Black hedge is at an advantage as it provides a perfect fit to VIX options by construction. In addition, if models are misspecified, allowing for changing structural model parameters may allow to capture structural shifts in the parameters and improve the hedging performance. In order to test these hypothesis, we also implement a calibration procedure to back out implied parameters and state variables from option quotes. Similar to Bakshi and Cao (2004), Broadie et al. (2007) and others, we minimize an implied volatility objective function as follows:

$$\min_{\Theta,\{x_1,\ldots,x_T\},\{m_1,\ldots,m_T\}} \frac{1}{\sum_t N_t^o}\sum_{t=1}^{T}\sum_{i=1}^{N_t^o}\big(IV(v_t,x_t,m_t,K_i,\tau_i,\omega_i) - IV^M(K_i,F_{t,\tau_i},\tau_i,\omega_i)\big)^2 + \frac{1}{\sum_t N_t^f}\sum_{t=1}^{T}\sum_{i=1}^{N_t^f} \times \big(F(v_t,x_t,m_t,\tau_i) - F^M(t,\tau_i)\big)^2. \qquad (9)$$

Several different implementations of this type of a procedure have been proposed in the literature. Bakshi et al. (1997) use option price data from one trading day, Bakshi and Cao (2004) use data collected during a full trading week, Broadie et al. (2007) use a range of randomized subsamples, whereas Christoffersen et al. (2009) alternate between optimizing over the structural parameters and the state variables. We follow Bakshi and Cao (2004) because their procedure appears a good compromise between frequent recalibration and imposing time-series stability in the parameters.[26] For the implementation of model-based hedge ratios, we use structural parameters that have been calibrated to options during the previous trading week. We then re-run the optimization in Eq. (9) fixing the parameter vector Θ to calibrate state variables

[26] In addition, we want to keep the hedging study purely our-of-sample, and hence some of the other calibration procedures such as Broadie et al. (2007) are less relevant in our context.

Table 5
In-sample Parameter Calibration. This table reports the parameter calibration results for all models defined in Eq. (1)–(3). The calibration period is from the last week of 2008 to December 2014 (the last week of 2008 is required to have out-of-sample parameters for the first week of 2009). The calibration is performed using minimum least square optimization on the objective function (9). Calibration are performed on weekly data and we report average calibrated parameters and their standard errors. Models are estimated to daily VIX option and VIX futures data described in Section 3.1.

Model	SVV	SVVC1	SVVC2	SVVJ	SVVJC1	SVVJC2
κ_v	20.4518	18.6428	19.1490	8.4189	8.9928	9.6348
	(0.4440)	(0.4838)	(0.4663)	(0.3721)	(0.2728)	(0.2793)
κ_x	4.0112	15.1544	15.6113	4.5796	11.2868	12.8524
	(0.2115)	(0.3361)	(0.3029)	(0.2861)	(0.3660)	(0.3426)
θ_x	4.7638	3.9085	3.7719	2.0076	1.6516	1.5993
	(0.1886)	(0.1266)	(0.1095)	(0.1310)	(0.0909)	(0.0691)
σ_x	11.5503	9.0364	9.3404	6.4344	4.9619	5.9512
	(0.0672)	(0.1328)	(0.1295)	(0.2317)	(0.1136)	(0.1392)
κ_m	3.2917	4.2805	3.6761	12.1342	8.7487	6.6244
	(0.3060)	(0.3116)	(0.2711)	(0.4425)	(0.3908)	(0.3341)
θ_m	3.7523	2.8774	2.8768	2.9091	2.9016	2.9062
	(0.0598)	(0.0371)	(0.0397)	(0.0283)	(0.0193)	(0.0218)
σ_m	0.5215	0.3884	0.5183	0.2410	0.4994	0.6886
	(0.0245)	(0.0238)	(0.0366)	(0.0210)	(0.0296)	(0.0373)
ρ_x		0.8513	0.8774		0.8593	0.8738
		(0.0128)	(0.0129)		(0.0101)	(0.0111)
ρ_m			−0.2112			−0.1917
			(0.0168)			(0.0148)
λ_v				1.4454	1.0654	1.6590
				(0.1071)	(0.0794)	(0.1353)
μ_J				0.3972	0.3761	0.3136
				(0.0128)	(0.0121)	(0.0123)
σ_J				0.2548	0.2563	0.2105
				(0.0065)	(0.0070)	(0.0074)
RMSE(F)	5.3847	2.1056	2.0151	3.3348	1.7280	1.5926
	(0.0736)	(0.0559)	(0.0574)	(0.0934)	(0.0469)	(0.0449)
RMSE(O)	3.6672	2.6096	2.5904	2.7416	2.3825	2.3677
	(0.0582)	(0.0370)	(0.0383)	(0.0352)	(0.0345)	(0.0349)

that minimize the pricing errors on the day the hedge is initiated. This approach guarantees that RMSHEs are purely out-of-sample.

We provide the estimated structural parameters in Table 5. As the model calibrations are based on data from the last week of December 2008 until the end of 2014, it is expected that the parameters will differ from the estimates in Table 2. In addition, it is well known in the equity index option literature, that some parameters show strong differences depending on the calibration method. For instance, Broadie et al. (2007) find that the time-series estimates for the vol-of-vol parameter in the Heston model (Heston, 1993) differs substantially from the parameters that are calibrated from options, albeit theoretically this parameter must be identical independent of the data source to rule out arbitrage opportunities. We find a similar pattern for σ_x where the Kalman filter estimations are substantially smaller than the calibrated parameters (the parameters are roughly twice as large in the calibration). A second structural difference in the parameters can be observed for the jumps, where the calibrations provide on average fewer jumps with λ_v dropping to values between 1.07 (SVVJC1) and 1.66 (SVVJC2). Given the lower frequency of jumps, on occurrence, their impact is much more pronounced with average jump sized between 0.31 (SVVJC2) and 0.40 (SVVJ). This result is also similar to equity index calibrations where the frequency of jumps varies significantly across estimation methods and sample periods (see for instance, Broadie et al., 2007; Duan & Yeh, 2010; Pan, 2002). Finally, our calibrations also find support for negative ρ_m values.

Table 6 provides hedging errors using the calibration parameters for the whole sample and various sub-categories. We focus on two main results. First, the performance differences using the calibrated parameters are much more pronounced, with clear improvements coming from both the inclusion of correlated state variables and jumps. Overall, the hedging errors drop from the SVV model (0.296) to almost half for SVVJC2 (0.164). Jump models now perform substantially better than before, indicating that jump parameters in particular may be affected by structural breaks over the sample period. Second, despite the more pronounced differences between models introduced in Section 2.1, all models still under-perform compared to the simple Black model hedge and the performance differences become even more pronounced with the Black hedge providing a RMSHE of almost half of the best performing structural model. These results are different from Alexander and Kaeck (2012) who find that re-calibration of a model may result in better hedging performance for FTSE 100 options. A possible explanation, explored further below, is that VIX option data are noisier than some of the more established index option markets and too frequent recalibration may result in over-fitting models.

4.4. Trader rules

Given the strong performance of the standard Black hedge ratio, we investigate if simple adjustments of the Black model result in further improvements in the hedging performance. We follow Derman (1999) who introduces three distinct volatility regimes and proposes adjustments of hedge ratios during the different market environments. The literature on hedge ratio adjustments for equity indices is comprehensive. Alexander, Rubinov, Kalepky, and Leontsinis (2012) provide an overview and employ a Markov-switching model to allow for time-variation in hedge ratio adjustments.

We now briefly review Derman's volatility regimes. The first regime, called *sticky strike*, assumes that the implied volatility remains constant with respect to the strike of the option, hence it coincides with using the Black model with the observed market implied volatility. Since market volatility is a function of strike and time-to-maturity, options with different strikes are hedged in separate trees. Derman suggests that such hedge ratios perform best during stable, trending equity markets. The *sticky moneyness* regime assumes that the implied volatility of an option changes

Table 6

Hedging RMSHE Calibration (daily rebalancing). This table reports the root-mean-square hedging error (RMSHE) of the hedging errors defined in Eq. (6) for a range of different hedging strategies and different option categories. Column 1 classifies the options depending on their moneyness level, where moneyeness is defined as K/F. Column 2 classifies options according to their remaining time-to maturity in days (*DTM*: days to maturity). *Num* in column 3 provides the number of options in the option category. Column 4 provides the RMSHE for an unhedged portfolio (*Unh*) for which the delta is set to zero. Column 5 provides the RMSHE for the Black model and columns 6 to 11 for the models introduced in Section 2.1. The hedging errors calculated based on the whole out-of-sample period from January 2009 until December 2014. The re-balancing frequency is daily.

Moneyness	DTM	Num	Unh	Black	SVV	SVVC1	SVVC2	SVVJ	SVVJC1	SVVJC2
All	All	119474	0.444	0.085	0.296	0.203	0.175	0.223	0.171	0.164
0.4–0.7	All	3421	0.674	0.074	0.281	0.152	0.146	0.286	0.126	0.114
0.7–0.9	All	30474	0.508	0.078	0.365	0.178	0.154	0.223	0.143	0.133
0.9–1.1	All	37712	0.430	0.093	0.329	0.223	0.189	0.238	0.186	0.180
1.1–1.3	All	24372	0.424	0.093	0.244	0.226	0.196	0.220	0.200	0.191
1.3–1.6	All	23495	0.348	0.074	0.160	0.179	0.156	0.186	0.153	0.146
All	≤ 90	76828	0.509	0.093	0.343	0.214	0.188	0.206	0.185	0.182
0.4–0.7	≤ 90	1375	0.932	0.065	0.385	0.164	0.168	0.171	0.122	0.112
0.7–0.9	≤ 90	17808	0.606	0.083	0.445	0.188	0.163	0.202	0.144	0.141
0.9–1.1	≤ 90	25151	0.479	0.101	0.377	0.232	0.200	0.230	0.198	0.197
1.1–1.3	≤ 90	17460	0.468	0.100	0.267	0.233	0.208	0.213	0.214	0.210
1.3–1.6	≤ 90	15034	0.412	0.083	0.179	0.190	0.171	0.159	0.171	0.167
All	≥ 91	42646	0.295	0.070	0.184	0.182	0.147	0.249	0.144	0.124
0.4–0.7	≥ 91	2046	0.420	0.079	0.179	0.143	0.130	0.342	0.128	0.115
0.7–0.9	≥ 91	12666	0.322	0.071	0.204	0.164	0.139	0.249	0.142	0.121
0.9–1.1	≥ 91	12561	0.309	0.073	0.202	0.204	0.163	0.254	0.158	0.141
1.1–1.3	≥ 91	6912	0.281	0.072	0.174	0.208	0.161	0.235	0.158	0.131
1.3–1.6	≥ 91	8461	0.190	0.056	0.119	0.157	0.125	0.227	0.114	0.097

whenever the underlying changes, but in a way that the implied volatility remains constant with respect to moneyness. If the implied volatility curve is upward sloping, then an increase in the futures price (and hence a decrease in the moneyness) results in a lower implied volatility. Due to this co-movement, the hedge ratio in the sticky moneyness regime differs from the sticky strike assumption for which changes in the underlying price process have no effect on implied volatilities. And third, local volatility models assume that the volatility moves in a deterministic fashion, depending on time and the underlying price, see Dupire (1994) and Derman and Kani (1994). This assumption is called *sticky tree*. Empirical and theoretical work investigating the implementation and performance of such hedge ratios include Coleman, Kim, Li, and Verma (2001) and Crépey (2004). Examples for applications of these rules to equity index data include Alexander et al. (2012), Alexander and Kaeck (2012), or Liu et al. (2019).

We implement the rules (which in our case translate into different vol-of-vol regimes) following Engelmann, Fengler, and Schwendner (2009). First, sticky strike deltas are calculated as before (as they are identical to the Black delta) and are labeled $\Delta^{SK}(K_i, \tau_i, \omega_i)$. We then estimate on each day t in our sample a simple regression as follows:

$$IV^{ma}(K_i, \tau_i, F_{t,\tau_i}, \omega_i) = \alpha_0 + \alpha_1 m_{i,t} + \alpha_2 m_{i,t}^2 + \alpha_3 m_{i,t}^3 + \varepsilon_{i,t} \tag{10}$$

where $m_{i,t} = \ln(K/F_{t,\tau_i})/\sqrt{\tau_i}$. Given the parameters of the regression model, it is straight-forward to calculate sticky moneyness hedge ratios by applying simple finite differences (for details we refer to (Engelmann et al., 2009)). As in Alexander and Kaeck (2012), we approximate the sticky tree regime by the following hedge ratio

$$\Delta^{ST}(K_i, \tau_i, \omega_i) = \Delta^{SK}(K_i, \tau_i, \omega_i) + \mathcal{V}^{SK}(K_i, \tau_i, \omega_i) \times \frac{\partial IV_t(K_i, \tau_i, F_{t,\tau_i}, \omega_i)}{\partial m_{i,t}} \frac{\partial m_{i,t}}{\partial K_i}$$

where $\mathcal{V}^{SK}(K_i, \tau_i, \omega_i)$ denotes the sticky-tree vega, i.e. the vega in the Black model using market implied volatility. Note that the slope of the smile term in this expression can also be easily estimated from the parametric regression model in Eq. (10).

Fig. 3 provides an example of the adjustments for a random day in our sample. It is apparent that the sticky tree and the sticky moneyness assumptions lead to corrections of opposite signs. While the sticky tree delta is larger than the Black delta, the sticky moneyness delta provides a lower hedge ratio due to the typical shape of the VIX option implied volatility functions. Also note that the adjustment is stronger for options with high moneyness levels and less pronounced (indeed vanishing) for options with low strikes.

Empirical hedging results for a daily rebalancing frequency for the various Black model adjustments are presented in Table 7. Over the whole sample period from 2009 until 2014 and for all options, we obtain a RMSHE of 0.096 for the sticky tree model and a RMSHE of 0.134 for the sticky moneyness assumption. Similar to the results for our model-based hedges of Section 4.1, our findings support that the higher hedge ratios implied by the sticky tree assumption lead to a deterioration of the hedging performance compared to the Black model which provides a RMSHE of 0.085. In addition, downward adjustments of the sticky moneyness model also lead to increased uncertainty in the hedging errors, such that overall the simple Black model outperforms various adjustments. Table 7 also provides hedging errors for other option categories, providing detailed results for various different moneyness/maturity combinations. The strong performance of the Black hedge is confirmed, as it remains the best performing model in all categories.

As the trading rules are designed to capture distinct market behavior it is particularly important to study whether sticky tree or sticky moneyness assumptions are performing better than the Black hedge during subsamples of our out-of-sample period. To this end, in the left panel of Fig. 4 we present RHSHEs for a 6-month sample period rolled over every trading day. Interestingly, the performance of the different hedging strategies is very consistent over time, with no clear evidence that one of the hedge ratio adjustments works particularly well during a certain market regime. The Black hedge remains the best hedging strategy throughout, the sticky tree model exhibits a comparable performance during the beginning and the end of the sample, and itself out-performs the sticky moneyness assumption consistently. Overall, we find no evidence that VIX options hedging is improved by considering distinct market regimes and an adjustment to Black hedge ratios. The weekly results in Table 8 and the right panel of Fig. 4 confirm these findings. In the right panel of Fig. 4 again the

Table 7
Hedging RMSHE for Trader Rules (daily rebalancing). This table reports the root-mean-square hedging error (RMSHE) of the hedging errors defined in Eq. (6) for a range of different hedging strategies and different option categories. Column 1 classifies the options depending on their moneyness level, where moneyness is defined as K/F. Column 2 classifies options according to their remaining time-to maturity in days (*DTM*: days to maturity). *Num* in column 3 provides the number of options in the option category. Column 4 provides the RMSHE for an unhedged portfolio (*Unh*) for which the delta is set to zero. Column 5 provides the RMSHE for the Sticky Tree model, column 6 for the sticky moneyness model and column 7 for the Black model. The hedging errors calculated based on the whole out-of-sample period from January 2009 until December 2014. The re-balancing frequency is daily.

Moneyness	DTM	Num	Unhedged	Sticky Tree	Sticky Moneyn.	Black (Sticky Strike)
All	All	119474	0.444	0.096	0.134	0.085
0.4-0.7	All	3421	0.674	0.109	0.084	0.074
0.7–0.9	All	30474	0.508	0.088	0.086	0.078
0.9–1.1	All	37712	0.430	0.099	0.133	0.093
1.1–1.3	All	24372	0.424	0.109	0.179	0.093
1.3–1.6	All	23495	0.348	0.085	0.140	0.074
All	≤ 90	76828	0.509	0.107	0.154	0.093
0.4–0.7	≤ 90	1375	0.932	0.139	0.082	0.065
0.7–0.9	≤ 90	17808	0.606	0.097	0.087	0.083
0.9–1.1	≤ 90	25151	0.479	0.109	0.149	0.101
1.1–1.3	≤ 90	17460	0.468	0.120	0.201	0.100
1.3–1.6	≤ 90	15034	0.412	0.097	0.164	0.083
All	≥ 91	42646	0.295	0.073	0.089	0.070
0.4–0.7	≥ 91	2046	0.420	0.083	0.085	0.079
0.7–0.9	≥ 91	12666	0.322	0.072	0.084	0.071
0.9–1.1	≥ 91	12561	0.309	0.077	0.094	0.073
1.1–1.3	≥ 91	6912	0.281	0.075	0.101	0.072
1.3–1.6	≥ 91	8461	0.190	0.059	0.079	0.056

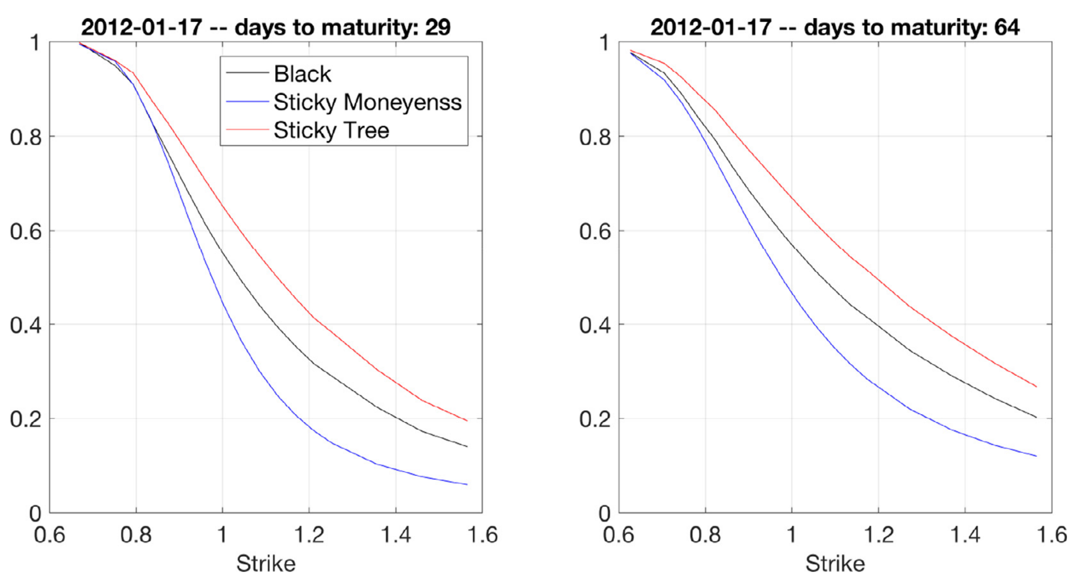

Fig. 3. Hedge Ratio Adjustments. This figure shows delta hedge ratios for the Black model and the sticky strike and stick moneyness assumption for VIX call options traded on January 17, 2012.

adjustments of the Black hedge strategy show roughly the same performance as the Black hedge, but are still consistently outperformed by it over the whole sample period. This results can also be seen in Column 7 of Table 8. The RMSHE presented here for the Black hedge strategy is smallest compared to sticky tree, sticky moneyness and Unhedged trading strategies over all moneyness levels and time to maturity groups.

4.5. Why does the Black delta perform so well?

Branger et al. (2012) show that even in controlled simulation exercises (which give researchers the advantage of knowing the exact model parameters and the level of latent state variables, and hence the true hedge ratios), the ad-hoc Black delta performs almost as well in reducing hedge error risk than the minimum-variance delta from the true data generating process (which in their study includes stochastic volatility and jumps). Branger et al. (2012) conclude that the Black delta, albeit being a partial derivative and not a minimum variance delta, is biased towards the true minimum variance hedge ratio. It is plausible that given the noise in real-world option data and the inherent estimation risk, model-based hedge ratios in our empirical setup lead to a deterioration in the hedge.[27]

To explore further the robustness of our Black hedge, we run several additional tests. Due to space constraints, we only provide a summary here and refer to Appendix A.7 for a more detailed description of these findings. As in Branger et al. (2012) we first

[27] Estimation risk has been found to work against sophisticated models in other empirical exercises. For instance, DeMiguel, Garlappi, and Uppal (2007) find that naive portfolio rules often outperform sophisticated models that are estimated with error.

Table 8
Hedging RMSHE for Trader Rules (weekly rebalancing). This table reports the root-mean-square hedging error (RMSHE) of the hedging errors defined in Eq. (6) for a range of different hedging strategies and different option categories. Column 1 classifies the options depending on their moneyness level, where moneyness is defined as K/F. Column 2 classifies options according to their remaining time-to maturity in days (*DTM*: days to maturity). *Num* in column 3 provides the number of options in the option category. Column 4 provides the RMSHE for an unhedged portfolio (*Unh*) for which the delta is set to zero. Column 5 provides the RMSHE for the Sticky Tree model, column 6 for the sticky moneyness model and column 7 for the Black model. The hedging errors calculated based on the whole out-of-sample period from January 2009 until December 2014. The re-balancing frequency is weekly.

Moneyness	DTM	Num	Unhedged	Sticky Tree	Sticky Moneyn.	Black (Sticky Strike)
All	All	22336	0.865	0.201	0.246	0.175
0.4–0.7	All	622	1.320	0.117	0.110	0.101
0.7–0.9	All	6016	1.009	0.159	0.175	0.150
0.9–1.1	All	7119	0.841	0.220	0.261	0.197
1.1–1.3	All	4571	0.799	0.237	0.300	0.190
1.3–1.6	All	4008	0.618	0.185	0.258	0.158
All	≤ 90	13411	0.982	0.231	0.285	0.201
0.4–0.7	≤ 90	243	1.778	0.139	0.109	0.105
0.7–0.9	≤ 90	3292	1.187	0.186	0.195	0.176
0.9–1.1	≤ 90	4434	0.939	0.250	0.296	0.226
1.1–1.3	≤ 90	3074	0.876	0.263	0.337	0.210
1.3–1.6	≤ 90	2368	0.731	0.214	0.305	0.181
All	≥ 91	8925	0.649	0.143	0.172	0.126
0.4– 0.7	≥ 91	379	0.911	0.100	0.110	0.099
0.7–0.9	≥ 91	2724	0.739	0.120	0.148	0.112
0.9–1.1	≥ 91	2685	0.650	0.156	0.187	0.137
1.1–1.3	≥ 91	1497	0.609	0.173	0.202	0.143
1.3–1.6	≥ 91	1640	0.403	0.132	0.167	0.118

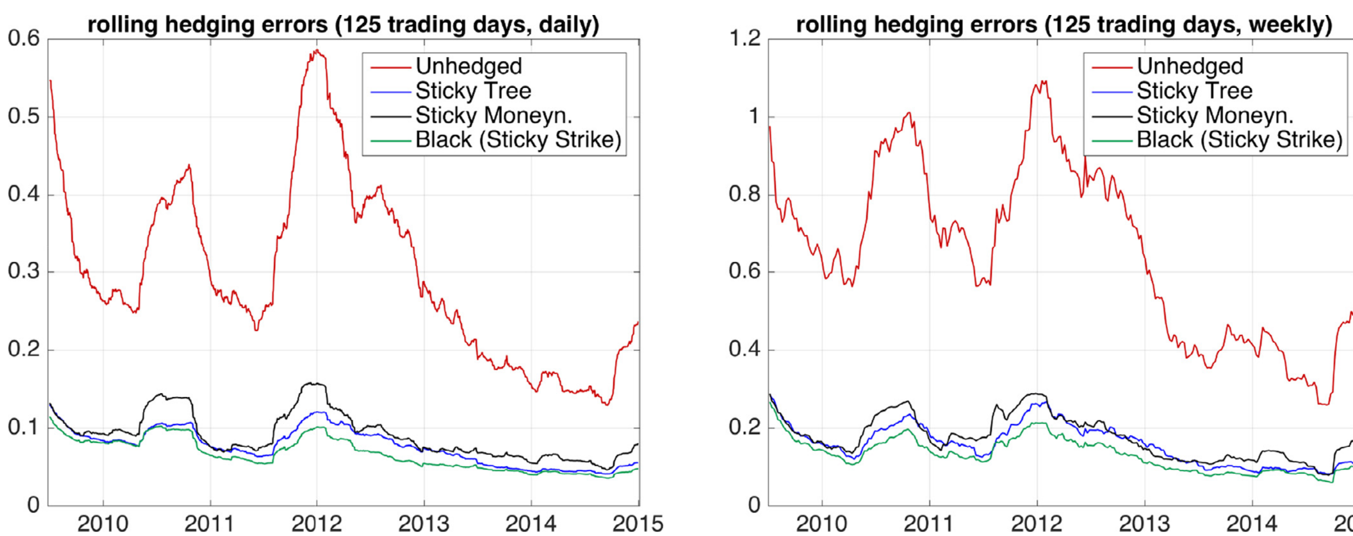

Fig. 4. Hedging Errors for Sub-Periods (Trader rules). This figure shows the average root-mean-square error (RMSHE) of the hedging errors defined in Eq. (6) for a range of different trader rules used for hedging. The RMSHE are calculated as follows. Left panel: For all hedging strategies we calculate the RMSHE using all hedging errors in the out-of-sample period (aggregating over different strikes and different maturities). We then average the RMSHE over the last 6 months and roll over on the next trading day. Right panel: for all hedging strategies we calculate the RMSHE using all hedging errors on every Wednesday in the out-of-sample period (aggregating over different strikes and different maturities). We then average the RMSHE over the last 6 months and roll over on the next Wednesday. *Unhedged* is the hedging strategy for which the delta is set to zero, Sticky Tree, Sticky Moneyness and Black-delta hedge are introduced in Section 4.4.

resort to a simulation exercise which allows us to control the *true* model. Our main focus in the simulation exercise is on the effect of noise in market data on the hedging performance. To this end, we add to the true simulated model an observation error, and vary the assumption on the volatility of the observation error and on its autocorrelation structure. Our simulation exercise confirms that model fit (and hence pricing error analysis) may not necessarily lead to better hedging performance. Indeed, in simulated markets the true model may be systematically outperformed by the Black hedge when option data is noisy and when pricing errors are autocorrelated. Building on these simulation results we then provide additional empirical hedging exercises to explore the extent to which pricing errors may affect our empirical hedging results. In our first exercise we explore how much of the superior hedge performance of the Black model is attributed to the fact that Black hedge ratios are calculated based on a perfectly fitted (albeit misspecified) model. More sophisticated models on the other hand will all produce smooth implied volatility curves and hence are not able to fit all market data perfectly (as in our simulation study). We find that about 30% of the inferior hedging performance is explained by model fit and conclude that a significant part of the superior performance of the Black model is related to the noise in the data. The second reason for the superior hedging performance of the Black model in our simulation exercise is that pricing errors are autocorrelated. To check whether this may indeed be the case for our option pricing models we estimate a simple AR(1) specification as in our simulation exercise and find strong support for this conjecture. Guided by the results of our simulation exercise, the empirical autocorrelation found in the errors of option pricing models as well as noisy data can explain why sophisticated model underperform in hedging exercises.

In Appendix A.7 we also explore whether it is possible at all in our sample to improve on the simple Black hedge. To this end, we calculate the ex-post optimal hedge ratio for each observation by simply choosing the delta that would have led to a zero hedging error. We average these ex-post deltas over the sample period across a fine grid of option moneyness buckets and compare these to (ex-ante) hedge ratios in alternative model specifications. This ex-post hedge ratio shows that the deltas of the Black model are much closer to this benchmark, in particular for at-the-money options for which the ex-ante Black model and the average optimal hedge are very close. The SVV model performs more poorly, especially around ATM. These additional results suggest that any parametric model, no matter how sophisticated, will struggle to improve on the Black hedge as this is already very close to the empirical ex-post optimum.

The superiority of the Black hedge may also be linked to potential overfitting. For instance, it is well established in the literature that supply and demand is an important determinant of option prices, see Garleanu, Pedersen, and Poteshman (2009). Any feature outside the modeling framework may lead to erroneous adjustments of hedge ratios as any model calibration/estimation necessarily translates effects outside the model into hedge ratio adjustments. Hentschel (2003) argues that measurement errors can also lead to systematic biases in volatility smiles. The comparison between our hedge strategies that are based on out-of-sample estimation and frequent recalibration support that estimation risk is a likely explanation for our results.

VIX option prices may also be partly driven by the fact that many market participants use simple Black hedge ratios in their daily risk management.[28] Under these assumptions, the demand and supply of options is a function of the hedge ratios calculated from the simple Black model, which may add to the success of the empirical hedging performance of the Black model we report in this paper.

4.6. Further results

In this section we comment on a number of additional results. We confine our discussion to the main findings to economize on space. All additional figures are provided in the Appendix. First, we repeat our hedging exercise with different hedging objective functions. In particular, we use the mean absolute hedging error (MAHE) and the relative version of RMSHE, which is based on the hedging error scaled by the price of the option, denoted C^M:[29]

$$\mathrm{MAHE}_t = \sqrt{\frac{1}{N_t^o}\sum_{i=1}^{N_t^o} |\ \mathrm{HE}_{i,t}|},$$

$$\mathrm{RMSRHE}_t = \sqrt{\frac{1}{N_t^o}\sum_{i=1}^{N_t^o}\left(\frac{\mathrm{HE}_{i,t}}{C_{it}^M}\right)^2}.$$

Our empirical findings for MAHE, which is used in Bakshi et al. (1997), are virtually indistinguishable from the results presented in the paper, and hence we do not report further details. Results for RMSRHE are summarized in Fig. A.1 and these findings also confirm the superiority of the Black hedge for a relative option hedging error objective function. Both, the overall model ranking and the robustness over time, is very similar to what was reported for RMSHE, and in particular none of our main conclusions change. The only notable difference using the relative hedge error objective function is that at times the overfitted jump model does not substantially improve on the unhedged position.

Our second set of results focuses on pricing errors. While the main contribution of this paper is to study hedging performance, in Fig. A.2 we provide further evidence regarding the pricing errors of our benchmark models. We find that the out-of-sample performance, using the pricing errors, yields a clear improvement when jumps and correlated state variables are considered (in line with Park, 2016). These additional results highlight the stark contrast between pricing and hedging error conclusions. The recalibration of models and a better fit to option data lead to clear out-of-sample improvements for option pricing exercises, our model hedging results are at odds with these pricing error results.

5. Conclusion

In this paper, we study the hedging performance of alternative VIX option pricing models and compare them with a range of simple hedging strategies that build on the Black model (Black, 1976). We find that simple models provide the best overall hedging performance. Adjusting hedging strategies for possible jumps in variance compared to a pure stochastic volatility setup leads to no significant performance improvement. We also show that a simple Black hedging strategy (assuming constant volatility) outperforms more elaborate structural models that consider stochastic volatility, long-term mean-reversion level, correlation structures between state variables and jumps. This result is robust for using different calibration methods, testing models on different subsamples, and for dividing the option sample in different moneyeness level and time to maturity groups.

Our results contradict previous evidence from the equity index market, where mixed results regarding the success of the Black-Scholes model hedge is found. Bakshi et al. (1997) find considerable improvements of stochastic volatility models when using a single implied-volatility parameter for all options, whereas Alexander and Kaeck (2012) find improvements when hedging is performed with each options implied volatility. Future research may provide further evidence into the success of advanced option pricing models by modeling VIX futures dynamics directly. This may improve the success of advanced models by assumption consistent with the price of the hedging instrument.

Supplementary material

Supplementary material associated with this article can be found, in the online version, at doi:10.1016/j.ejor.2019.11.034.

References

Alexander, C., & Kaeck, A. (2012). Does model fit matter for hedging? Evidence from FTSE 100 options. *Journal of Futures Markets, 32*(7), 609–638.
Alexander, C., & Korovilas, D. (2013). Volatility exchange-traded notes: Curse or cure? *The Journal of Alternative Investments, 16*(2), 52–70.
Alexander, C., & Nogueira, L. M. (2007). Model-free hedge ratios and scale-invariant models. *Journal of Banking and Finance, 31*(6), 1839–1861.
Alexander, C., Rubinov, A., Kalepky, M., & Leontsinis, S. (2012). Regime-dependent smile-adjusted delta hedging. *Journal of Futures Markets, 32*(3), 203–229.
Andreou, P. C., Charalambous, C., & Martzoukos, S. H. (2008). Pricing and trading European options by combining artificial neural networks and parametric models with implied parameters. *European Journal of Operational Research, 185*, 1415–1433. doi:10.1016/j.ejor.2005.03.081.
Bakshi, G., & Cao, C. (2004). Disentangling the contribution of return-jumps and volatility jumps: Insights from individual equity options. *Working paper, University of Maryland*.
Bakshi, G., Cao, C., & Chen, Z. (1997). Empirical performance of alternative option pricing models. *Journal of Finance, 52*(5), 2003–2049.
Ballotta, L., Deelstra, G., & Rayée, G. (2017). Multivariate FX models with jumps: Triangles, Quantos and implied correlation. *European Journal of Operational Research, 260*(3), 1181–1199.

[28] We would like to thank an anonymous referee for suggesting this.
[29] Hedging errors for put options are defined accordingly.

Bandi, C., & Bertsimas, D. (2014). Robust option pricing. *European Journal of Operational Research, 239*(3), 842–853. doi:10.1016/j.ejor.2014.06.002.

Bao, Q., Li, S., & Gong, D. (2012). Pricing VXX option with default risk and positive volatility skew. *European Journal of Operational Research, 223*(1), 246–255. doi:10.1016/j.ejor.2012.06.006.

Bardgett, C., Gourier, E., & Leippold, M. (2019). Inferring volatility dynamics and risk premia from the S&P 500 and VIX markets. *Journal of Financial Economics, 131*(3), 593–618. doi:10.2139/ssrn.2296826.

Bates, D. S. (2000). Post-'87 crash fears in the S&P 500 futures option market. *Journal of Econometrics, 94*(1–2), 181–238.

Bates, D. S. (2003). Empirical option pricing: A retrospection. *Journal of Econometrics, 116*(1–2), 387–404.

Bates, D. S. (2006). Maximum likelihood estimation of latent affine processes. *Review of Financial Studies, 19*(3), 909–965.

Berger, D., Dew-Becker, I., & Giglio, S. (2019). Uncertainty Shocks as Second-Moment News Shocks. *The Review of Economic Studies*. doi:10.1093/restud/rdz010.

Black, F. (1976). The pricing of commodity contracts. *Journal of Financial Economics, 3*(1–2), 167–179.

Black, F., & Scholes, M. S. (1973). The pricing of options and corporate liabilities. *Journal of Political Economy, 81*(3), 637–654.

Bloom, N. (2009). The impact of uncertainty shocks. *Econometrica, 77*(3), 623–685. doi:10.3982/ECTA6248.

Bondarenko, O. (2014). Why are put options so expensive? *The Quarterly Journal of Finance, 4*(3), 1450015.

Branger, N., Krautheim, E., Schlag, C., & Seeger, N. (2012). Hedging under model misspecification: All risk factors are equal, but some are more equal than others.... *Journal of Futures Markets, 32*(5), 397–430.

Broadie, M., Chernov, M., & Johannes, M. (2007). Model specification and risk premia: Evidence from futures options. *Journal of Finance, 62*(3), 1453–1490.

Broadie, M., & Detemple, J. B. (2004). Anniversary article: Option pricing: valuation models and applications. *Management Science, 50*(9), 1145–1177. doi:10.1287/mnsc.1040.0275.

Carr, P., Geman, H., Madan, D. B., & Yor, M. (2002). The fine structure of asset returns: An empirical investigation. *The Journal of Business, 75*(2), 305–333.

Carr, P., & Wu, L. (2007). Stochastic skew in currency options. *Journal of Financial Economics, 86*(1), 213–247. doi:10.1016/j.jfineco.2006.03.010.

Chockalingam, A., & Muthuraman, K. (2015). An approximate moving boundary method for American option pricing. *European Journal of Operational Research, 240*(2), 431–438. doi:10.1016/j.ejor.2014.07.031.

Christoffersen, P., Dorion, C., Jacobs, K., & Karoui, L. (2014). Nonlinear Kalman filtering in affine term structure models. *Management Science, 60*(9), 2248–2268.

Christoffersen, P., Heston, S., & Jacobs, K. (2009). The shape and term structure of the index option smirk: Why multifactor stochastic volatility models work so well. *Management Science, 55*(12), 1914–1932.

Christoffersen, P., & Jacobs, K. (2004). The importance of the loss function in option valuation. *Journal of Financial Economics, 72*(2), 291–318.

Christoffersen, P., Jacobs, K., & Mimouni, K. (2010). Volatility dynamics for the S&P 500: Evidence from realized volatility, daily returns, and option prices. *Review of Financial Studies, 23*(8), 3141–3189.

Christoffersen, P., Jacobs, K., & Ornthanalai, C. (2012). Dynamic jump intensities and risk premiums: Evidence from S&P 500 returns and options. *Journal of Financial Economics, 106*(3), 447–472.

Coleman, T., Kim, Y., Li, Y., & Verma, A. (2001). Dynamic hedging with a deterministic local volatility function model. *Journal of Risk, 4*(1), 63–89.

Coqueret, G., & Tavin, B. (2016). An investigation of model risk in a market with jumps and stochastic volatility. *European Journal of Operational Research, 253*(3), 648–658.

Corsaro, S., Kyriakou, I., Marazzina, D., & Marino, Z. (2019). A general framework for pricing Asian options under stochastic volatility on parallel architectures. *European Journal of Operational Research, 272*(3), 1082–1095. doi:10.1016/j.ejor.2018.07.017.

Crépey, S. (2004). Delta-hedging vega risk? *Quantitative Finance, 4*, 559–579.

Date, P., & Islyaev, S. (2015). A fast calibrating volatility model for option pricing. *European Journal of Operational Research, 243*(2), 599–606. doi:10.1016/j.ejor.2014.12.031.

DeMiguel, V., Garlappi, L., & Uppal, R. (2007). Optimal versus naive diversification: How inefficient is the 1/n portfolio strategy? *The Review of Financial studies, 22*(5), 1915–1953.

Dennis, P., & Mayhew, S. (2009). Microstructural biases in empirical tests of option pricing models. *Review of Derivatives Research, 12*(3), 169–191. doi:10.1007/s11147-009-9039-0.

Derman, E. (1999). Regimes in volatility. *Risk, 12*(4), 55–59.

Derman, E., & Kani, I. (1994). Riding on a smile. *Risk, 7*(2), 277–284.

Detemple, J. B., & Osakwe, C. (2000). The valuation of volatility options. *European Finance Review, 4*(1), 21–50.

Dotsis, G., Psychoyios, D., & Skiadopoulos, G. (2007). An empirical comparison of continuous-time models of implied volatility indices. *Journal of Banking and Finance, 31*(12), 3584–3603.

Duan, J.-C., & Yeh, C.-Y. (2010). Jump and volatility risk premiums implied by VIX. *Journal of Economic Dynamics and Control, 34*(11), 2232–2244.

Duffie, D., Pan, J., & Singleton, K. (2000). Transform analysis and asset pricing for affine jump-diffusions. *Econometrica, 68*(6), 1343–1376.

Dumas, B., Fleming, J., & Whaley, R. E. (1998). Implied volatility functions: Empirical tests. *Journal of Finance, 53*(6), 2059–2106.

Dupire, B. (1994). Pricing with a smile. *Risk, 7*(1), 18–20.

Egloff, D., Leippold, M., & Wu, L. (2010). Variance risk dynamics, variance risk premia, and optimal variance swap investments. *Journal of Financial and Quantitative Analysis, 45*(5), 1279–1310.

Engelmann, B., Fengler, M. R., & Schwendner, P. (2009). Hedging under alternative stickiness assumptions: An empirical analysis for barrier options. *The Journal of Risk, 12*(1), 53.

Eraker, B. (2004). Do stock prices and volatility jump? Reconciling evidence from spot and option prices. *Journal of Finance, 59*(3), 1367–1404.

Fabozzi, F. J., Paletta, T., Stanescu, S., & Tunaru, R. (2016). An improved method for pricing and hedging long dated American options. *European Journal of Operational Research, 254*(2), 656–666. doi:10.1016/j.ejor.2016.04.002.

Fabozzi, F. J., Paletta, T., & Tunaru, R. (2017). An improved least squares Monte Carlo valuation method based on heteroscedasticity. *European Journal of Operational Research, 263*(2), 698–706. doi:10.1016/j.ejor.2017.05.048.

Fu, M. C., Li, B., Li, G., & Wu, R. (2016). option pricing for a jump-diffusion model with general discrete jump-size distributions. *Management Science, 63*(11), 3961–3977. doi:10.1287/mnsc.2016.2522.

Fusai, G., Germano, G., & Marazzina, D. (2016). Spitzer identity, Wiener–Hopf factorization and pricing of discretely monitored exotic options. *European Journal of Operational Research, 251*(1), 124–134. doi:10.1016/j.ejor.2015.11.027.

Garleanu, N., Pedersen, L. H., & Poteshman, A. M. (2009). Demand-based option pricing. *Review of Financial Studies, 22*(10), 4259–4299.

Grunbichler, A., & Longstaff, F. A. (1996). Valuing futures and options on volatility. *Journal of Banking & Finance, 20*(6), 985–1001.

Hansen, P. R., Lunde, A., & Nason, J. M. (2011). The Model Confidence Set. *Econometrica, 79*(2), 453–497. doi:10.3982/ECTA5771.

Hentschel, L. (2003). Errors in implied volatility estimation. *Journal of Financial and Quantitative analysis, 38*(4), 779–810.

Heston, S. L. (1993). A closed-form solution for options with stochastic volatility with applications to bond and currency options. *Review of Financial Studies, 6*(2), 327–343.

Hirsa, A. (2016). *Computational methods in finance*. CRC Press.

Hull, J., & White, A. (2017). Optimal delta hedging for options. *Journal of Banking & Finance, 82*, 180–190.

Jackwerth, J. C., & Rubinstein, M. (1996). Recovering probability distributions from option prices. *The Journal of Finance, 51*(5), 1611–1631.

Jin, X., Li, X., Tan, H. H., & Wu, Z. (2013). A computationally efficient state-space partitioning approach to pricing high-dimensional American options via dimension reduction. *European Journal of Operational Research, 231*(2), 362–370. doi:10.1016/j.ejor.2013.05.035.

Julier, S. J., & Uhlmann, J. K. (1997). New extension of the Kalman filter to nonlinear systems. In *Aerosense'97* (pp. 182–193). International Society for Optics and Photonics.

Kaeck, A., & Alexander, C. (2013). Continuous-time VIX dynamics: On the role of stochastic volatility of volatility. *International Review of Financial Analysis, 28*, 46–56.

Kou, S. (2002). A jump-diffusion model for option pricing. *Management Science, 48*(8), 1086–1101.

Liu, X., Cao, Y., Ma, C., & Shen, L. (2019). Wavelet-based option pricing: An empirical study. *European Journal of Operational Research, 272*(3), 1132–1142. doi:10.1016/j.ejor.2018.07.025.

Mayhew, S. (2002). Competition, market structure, and bid-ask spreads in stock option markets. *Journal of Finance, 57*(2), 931–958. doi:10.1111/1540-6261.00447.

Mencía, J., & Sentana, E. (2013). Valuation of VIX derivatives. *Journal of Financial Economics, 108*(2), 367–391.

Mercuri, L., & Rroji, E. (2018). Option pricing in an exponential MixedTS Lévy process. *Annals of Operations Research, 260*(1–2), 353–374. doi:10.1007/s10479-016-2180-x.

Ornthanalai, C. (2014). Levy jump risk: Evidence from options and returns. *Journal of Financial Economics, 112*(1), 69–90.

Pan, J. (2002). The jump-risk premia implicit in options: Evidence from an integrated time-series study. *Journal of Financial Economics, 63*, 3–50.

Park, Y.-H. (2016). The effects of asymmetric volatility and jumps on the pricing of VIX derivatives. *Journal of Econometrics, 1*, 313–328.

Schönbucher, P. J. (1999). A market model for stochastic implied volatility. *Philosophical Transactions: Mathematical, Physical and Engineering Sciences, 357*(1758), 2071–2092. Mathematics of Finance

Wan, E., & Van Der Merwe, R. (2000). The unscented Kalman filter for nonlinear estimation. In *Adaptive systems for signal processing, communications, and control symposium 2000. AS-SPCC. the IEEE 2000* (pp. 153–158). IEEE.

Whaley, R. (1993). Derivatives on market volatility: Hedging tools long overdue. *Journal of Derivatives, 1*, 71–84.

Wong, H. Y., & Lo, Y. W. (2009). Option pricing with mean reversion and stochastic volatility. *European Journal of Operational Research, 197*(1), 179–187. doi:10.1016/j.ejor.2008.05.014.